

nature

12 May 1977

A modest proposal to the Department of Education and Science

THERE is by no means enough lobbying of government on behalf of science. That was the challenge thrown out last week by Mrs Shirley Williams, Britain's Secretary of State for Education and Science, at a gathering of the Association of British Science Writers. The Arts, for which she is also responsible, lobby regularly and aggressively; science's voice is rarely heard. This will come as a surprise to many scientists who have assumed that lobbying would be unwelcome and counter-productive. But the challenge is not just to scientists; it is to Mrs Williams herself to find ways of opening up the London machinery to the community at large so that access to decision-makers becomes easier to achieve.

That British scientists don't do enough to express their opinions is undeniable. Certainly when an issue like the payment of CERN's subscription comes up there is a great flurry of activity, but for the most part the British scientist is totally silent—content, presumably, with zero growth rates and a secure job. Maybe the assumption is that the media will look after all questions concerned with policy and will spot all issues which ought to be raised; and then, when there is nothing in *Nature*, *New Scientist* or the *Times Higher Education Supplement*, it is assumed that there is really nothing to worry about. This would be a dangerous assumption; the media cannot hunt out what it knows nothing about, and certainly in the past few years the number of individual scientists who have approached us with a request that we should look into something or pursue a particular campaign is pathetically small.

Or maybe it is widely believed that worthy bodies like the Advisory Board for the Research Councils or the individual research councils themselves adequately look after the interests of science. The trouble with that assumption is that these bodies already have a fairly clear and well-charted relationship with the Department of Education and Science, and this allows relatively limited scope for anything resembling lobbying, which would, in any case, have to be done by chairmen themselves. What,

presumably, Mrs Williams is looking for is a regular flow of opinion of diverse sorts, not just from research council heads, on how support for science should develop.

It would be too easy, however, to lay the blame for the lack of lobbying entirely on apathy in the country. That apathy has very largely been fostered by the way in which British society divides its institutions so scrupulously. There is Parliament. There is the civil service. There is the law. There is industry and commerce. There is the military. There is the academic world. And there is the press. All too rarely do people find it easy to cross these frontiers and live happily in two or more camps either concurrently or consecutively, and as a result public life suffers through narrowed horizons.

Obviously there is no simple remedy to this. Over-secure career prospects and innate beliefs that certain ways of earning a living are superior to others are not going to be done away with overnight. But there is some positive action Mrs Williams could take to get to hear the scientist's voice; she could create a post for a distinguished scientist who with a small staff could act as the department's ears and eyes into the scientific community—and not just academics. Maybe such a person would be called Chief Scientist, but his or her function would be rather different from those of chief scientists in other ministries. In many ways the job of being a link with practising scientists used to be performed by the government's chief scientific adviser, but now that this position has been allowed to lapse there is no one to whom scientists can turn in a personal capacity. The department's chief scientist ought to fulfil that role, but in doing so ought to be highly mobile. The attractions of life in London close to the corridors of power are obvious, but the right person would not be bound to an office but would be regularly on the road, thereby not just hearing what institutions and professional bodies have to say, but more importantly seeking out the opinions of individuals. □



Today's risks: thinking the unthinkable

Reg Farmer, Safety Adviser to the UK Atomic Energy Authority and Visiting Professor at Imperial College of Science and Technology, offers his perspective on risk in modern society

THERE is growing concern over the safety of our industrial and social activities. People ask, "Is it safe?" For many years we have been told about the risks from radiation; now we hear of the potential hazards to people or the environment from mercury, lead, asbestos, aerosols, and recently from saccharin.

There have been various attempts to draw a dividing line between 'safe' and 'unsafe' or even to establish absolute safety—a state of no risk. In the early days of atomic energy in the UK, there was an attempt to find a design of reactors which was inherently safe—any move to danger would be self-cancelling. More recently some people claim that if any risk can be identified, it must be eliminated. Neither of these approaches has proved to be achievable in practice.

In the USA it was first implied that reactors would be safe if contained, and there was a proposal to design a 'zero release' containment, so as to appear to give absolute safety. Later there grew a recognition of some risk—in the USA a three-man hearing was required to satisfy itself that the proposed nuclear plant created no undue risk. 'Undue' did not mean no risk, but it introduced the concept of reasonable. No specific directive was given as to that which might be judged reasonable other than through the circular statement:

To understand 'undue risk' in this context it is necessary to understand 'reasonable assurance' and AEC has defined this as meaning 'reasonable probability, establishing under the particular circumstances of the case in good faith and in the exercise of sound discretion and expert judgment.'

A not dissimilar situation arises in relation to the use of some chemicals, fertilisers, insecticides, food additives and so on. Generally their use should be controlled by the best possible means. In some cases and in some countries absolute criteria apply—as in the US Food and Drug Administration requirement that any foodstuff shown to produce cancer shall be banned. It is not then a question of what is reasonable, but a pursuit of no risk. This absolute criterion is not in keeping with the acceptance of risk at work, at play, or risk through the continued low level contamination of our environment.

It would be convenient if the answer to the question, "Is it safe?" could be yes or no. For thirty years, throughout the development of atomic energy, there have been many and varied attempts to derive criteria which, if satisfied, would imply safety or adequate safety and, if not satisfied, would imply some risk. For some years it was claimed that pressure vessels could not fail catastrophically if they were maintained at some appropriate temperature (nil-ductility temperature + 60 °F). There have been many attempts to find an all-embracing solution, by proposing an upper limit accident which the system should be designed to tolerate safely, to circumscribe all lesser accidents—such as the assumption that the largest pipe could fail or that, in gas-cooled reactors, the coolant flow might cease instantaneously.

These early accident studies followed a pattern, now called deterministic; this emerged when it was recognised that, in general, no system is inherently safe. It was at one time suggested that we should take account of two co-incidental faults, but not three. This approach has been

revised in various ways, such as a requirement that there should be two or three 'boundaries' between fission products and man. These requirements are largely illusory and do not define what risk level is being accepted.

I would like now to consider some other aspects of the current risk scene. It would appear that some hazards are thought to be more unacceptable than others, such as the results of thalidomide, the effect of dioxin or of radiation. It is easy to understand the horror of continued evidence of mistake or misfortune, as presented by malformed children. It is not easy to identify or justify the fear of exposure to radiation as against exposure to toxic chemicals.

Radiation and radioactive substances have been with us always; some say the new situation arises because of the bomb; others that it arises from the fear of cancer, as may be evidenced by the USA ruling on the prohibition of any potentially carcinogenic substances in food. If I were given the choice of death today by accident, or death in 10–20 years by cancer, which would I choose? There is some chance—about 1 in 5—of dying by cancer from so-called natural causes. Alternatively, and indeed more likely in an accident, how would I rate a chance of 1 in 10 to 1 in 100 of immediate death or death within a short time through explosions or toxic gases, as against the same risk of death in 10 to 20 years? I would choose to live on.

It is believed that radiation presents new risks, but radiation has always been with us, and carcinogenic or genetic damage does not arise solely from radiation, as reported by Sir Edward Pochin¹. Apart from 'a different sort of hazard' from radiation, it was thought that accidents to nuclear sites might cause greater harm than hitherto accepted in modern society. In my view this is not so, and at this point I approach the nub of this present discourse, that is, how best or in what way to describe the results of a hypothetical accident to an industrial site?

In my early days in nuclear energy, it was customary to talk about 'maximum credible accidents' and particularly so in the USA, but in general so to circumscribe the conditions of the hypothetical accident as to make the 'mca' acceptable, so that the installation might be licensed. However, some writers describe accidents beyond those currently reviewed in the traditional 'mca' approach, as in the early document, WASH-740, describing in 1957 the possibility of 3,400 deaths and more than 10,000 injuries. Somewhat earlier, in 1955, at Geneva official papers described hazards from radioactive substances as one million to one thousand million times more dangerous than chemical poisons, whereas it is now generally accepted that plutonium is only about ten times more toxic than chlorine. This violent see-saw, from the trivial to the disastrous, has dogged the development of atomic energy. We need to come to terms with this situation. Perhaps both ends of the scale are possible and both represent half-truths, but we may penalise the further development not only of energy resources but also of our industrial and technological development unless we have a better understanding of the nature of hazards and their associated risk rates.

An accident in any major industry may lead to financial loss, the shutdown of plant; it may lead to no harm, or the same accident might cause many casualties. There is some chance of an accident occurring in most industrial installations or in transport. The consequences will depend upon many things: the quantity of toxic or explosive material available; the quantity released in the accident; the rate of release; whether the release occurs in seconds, minutes or hours; the surrounding population distribution; the conditions of wind and weather; topography; the effectiveness of remedial procedures on and off site.

Some of these are reasonably independent or capable of prior assessment. In many cases the accident is not caused by wind or weather, but the consequences may be greatly influenced by them. In some cases accidents are caused and may be predicted from unusual weather patterns—flood, high wind, low temperature and so on. In general, there is no clearly defined maximum; it is possible to derive more and more casualties by presupposing more unusual conditions—the limit or maximum remains vague and unapproachable. Yet this idea of 'worst' or 'maximum' cannot be too lightly discarded. In financial matters it is usual to have a defined limit of liability; by the same token it would be desirable to have an upper limit of hazard which may then be accepted or rejected.

Unfortunately this route is not yet open to us, and is certainly one not easy to follow. In recent history it has been the practice to describe the worst accident to reactors or to some industrial plants in terms of some framework of credibility, generally so framed as to make the consequences of the accident relatively small and not too difficult to accept.

There is another route—an alternative to the synthetic analysis of a hypothetical accident. This route is to list and survey the accidents which have occurred in certain fields—ground or air transport, petrochemicals, the manufacture of chlorine and so on—and to show that from some dozen or so accidents only a few people are killed. It is then fair to extrapolate to the future if the analysis takes account of any change in toxicity and quantity of material at risk and any improvement in the means for their control. Indeed, estimates prepared in this way give a good indication of the 'likely' results of an accident.

To return to 'the worst'. Until recently it might 'reasonably' have been said that the worst aircraft accident might kill 350 people. There was always some risk that two might collide and this has happened. It could be worse. USA studies reported in evidence to a Senate Committee⁷ have given a probability of an aircraft crash on Hollywood Park grandstand whilst occupied by a large crowd as about 10^{-5} per year and the chance of killing 100–10,000 as being in the range 10^{-5} – 10^{-7} per year. In this simple example, it could be that of two groups discussing risks in air transport, one could believe that 350 casualties represented the 'worst' whereas the other could fear the risk of 10,000 casualties.

An even greater range could apply to some industrial accidents. The accidental release of toxic or explosive gases has caused 10–20 deaths on several occasions. A retrospective analysis shows that in some cases, had the wind been blowing in another direction, the result might have been ten times worse. Furthermore, had the weather pattern been less favourable, the casualties would have been greater. It is known that under inversion conditions, with low wind speed, the downwind air concentration for a given release of material is up to ten times higher at a given distance than for average weather conditions, and that in the UK, inversion occurs about 15% of the time—a small chance, but still significant.

There has been considerable effort employed in assessing the chance of major hazards from man-made, or natural disasters, as from explosion, fire, earthquake, or even meteorite impact. These studies require a combination of two factors: extrapolation from past history weighted according to the judgment of the assessor; and the use of currently available data on probability of wind direction, weather category, population distribution.

Many small accidents have occurred, and will continue to occur, which kill from 1–10 people. There is some small chance that a combination of circumstances might cause the death of over 100 people. Rasmussen² and Okrent³ have explored a range of possible hazards in the USA and, for each of various categories of fire, explosion and so on, have given their prediction as follows:

Chance of killing more than 100 people:	10^{-1} – 10^{-2} per year
Chance of killing more than 1,000 people:	10^{-2} – 10^{-4} per year
For earthquakes, they estimate:	
Chance of killing more than 10,000 people:	$\sim 10^{-2}$ per year
And for dam failures:	
Chance of killing more than 10,000 people:	$\sim 10^{-3}$ per year
Their assessment for overall risk from all recognised sources gives:	
Chance of killing more than 1,000 people:	10^{-1} per year
Chance of killing more than 10,000 people:	in the range 10^{-2} – 10^{-3} per year

The situation in the USA is different from that in the UK. Natural hazards from tornadoes, dams and earthquakes are greater in the USA. On the other hand the UK population density is roughly 10 times that of the USA. My own view is that the comparable risk in the UK lies within a factor of 10 of the USA—perhaps a factor of 10 lower. But on this matter we should not be too complacent.

These figures may be frightening, and it would be terrible if 1,000 to 10,000 people were killed; it is unlikely, but it could happen and is experienced in other countries from earthquakes and floods. Even so, a single event of this type would barely change the overall accidental death rate in the USA or Europe. The overall accidental death rate is about 1 in 2,500 per year, although this is age dependent. Most of us have a 2% chance of accidental death in 50 years. In a country of 50 million people, this means that more than 20,000 die from theoretically avoidable accidental death (including suicides) each year.

In giving guidance on some new activity, there is need to take account of both current risk levels and public reaction but, in so doing, it is important that the public should be informed and that they should recognise the need for wise and balanced judgment in situations where absolute safety can no longer exist. The public seem to expect the experts to seek ways of reducing risk and if ways are discovered which promise clear benefit at not too great cost, they expect the remedy to be applied.

Additionally, I believe that there are pressures to reduce risk irrespective of the risk rate, or in the absence of specific knowledge of the risk rate. There will be pressure to reduce risks in atomic energy, in aircraft flight and ground transport, and in many industries, such as mining, fishing, chemicals, petrochemicals, steel, oil exploration; and also pressure to spend more money in medical research and health facilities as a means of reducing risk of suffering or death.

How best to resolve these pressures on money and skill? How best to decide whether to reduce risks still lower in places where they are already low, as against the need to reduce the frequent and relatively random causes of common accidents? There is no easy answer.

I believe that progress should be made in two directions: to improve communication between people on this matter of risk of hazards to life and the environment; and to continue to improve the techniques for the assessment of the risk of future harm to man and environment. A considerable effort is being deployed in both directions in the atomic energy field and is now being extended more widely in our industrial activities. □

¹ Pochin, E. E. *Estimated Population Exposure from Nuclear Power Production and Other Radiation Sources*, NEA/OECD, 1976.

² WASH-1400, *Reactor Safety Study—An Assessment of Accident Risks in US Commercial Nuclear Power Plants*, USNRC, October 1975.

³ Okrent, D. Investigation of Charges relating to Nuclear Reactor Safety, Hearing before the Joint Committee on Atomic Energy, Congress of the United States, 94 Congress, Second Session, 2, Appendixes 12–19.

Zoo story

Alastair Hay looks at the London Zoo, now experiencing some changes at the top

AT its Annual General Meeting this week, the Zoological Society of London (ZSL) was due to witness a major change in the composition of the Council, its governing body. Both the President, Prince Philip, and the Secretary, Lord Zuckerman, were resigning their offices, with Prince Philip leaving the Council altogether. The Council nominated Lord Zuckerman as President and Dr Ronald Hedley, the Director of the British Museum (Natural History), for the Secretaryship. The Secretary of the ZSL is the executive officer of the Council; when it was a salaried post in the years 1847–1943 it was one of the most sought-after positions by zoologists in Britain.

Three years after its inception in 1826—under the auspices of Sir Stamford Raffles (the first President) and the then President of the Royal Society Sir Humphrey Davy—the ZSL was incorporated by Royal Charter with the specific objects of “ensuring the advancement of zoology and animal physiology and the introduction of new and curious subjects of the animal kingdom”. Animal physiology at this time was concerned primarily with the optimal conditions for animal breeding. Thus, whereas Raffles was more interested in the scientific potential of the Zoo, Davy—in Raffles’ view—was more concerned with its “practical and immediate utility to the country gentlemen”.

Davy’s ideas notwithstanding, his support enabled the ZSL to develop rapidly into the centre for zoological discussion in Britain. In previous years the Linnean Society had been the hub of these discussions, although it was principally concerned with botany. There is some irony in the fact that the foundation of the ZSL was a reaction against this bias; the Linnean’s own development, in 1788 to cultivate “the science of natural history in all its branches”, was itself a reaction to the Royal Society’s own preoccupation with the physical sciences.

During its first hundred years the ZSL’s publications were mainly concerned with systematics, comparative anatomy, taxonomy and phylogeny. In this respect they very much reflected the interest of the Secretaries of the Society, men such as Philip Lutley Sclater (1859–1902) and Peter Chalmers Mitchell (1902–1935). But well before Sclater retired, the biological sciences were themselves undergoing rapid change. Zoology and animal phy-

siology were now sciences which required animal experimentation and a knowledge of genetics, biochemistry and endocrinology. The Establishment at the Zoo seemed to be unaware of these developments and consequently younger zoologists were looking elsewhere. The Society was losing the scientific support on which its integrity and very survival depended.

Huxley’s problems

Julian Huxley was the first of the experimental zoologists to be appointed Secretary. Although he attempted to change the Society’s approach to zoology he was, in the main, unsuccessful. By 1942 Huxley and the council were openly opposed to one another. Huxley resigned in 1943, an action which was regarded by many zoologists as a serious defeat for progressive scientists. For the next twelve years few changes were possible in the Society. The situation for the duration of the war and the first decade of post-war Britain required Huxley’s successors, Sheffield Neave and Lord Chaplin, to devote all available resources to restoration following bomb-damage and to the replenishment of the animal collections. Lord Zuckerman became Secretary in 1955 with the ZSL in a very poor state.

The ZSL received confirmation of its status as an educational charity in 1930. Lord Zuckerman believes implicitly in the importance of the Zoo as an educational institution and in the need for the Society to adhere to the terms of its Charter and his determination to institute changes along these lines and to bring it back into the mainstream of zoology involved him in a major clash with Fellows of the Society in 1957. At the Annual General Meeting that year he proposed, on behalf of the Council, an increase in Fellows’ subscriptions (unchanged for over 100 years), the creation of a new class of Scientific Fellow, and the extension of public visiting hours on Sundays.

The changes were essential if the Society was to preserve its charitable status and to increase its revenue. For many of the Fellows, however, these proposals represented a considerable challenge to their privileged status. As a result the 1957 meeting was probably one of the most lively in the Society’s history. The Council’s proposals were finally ratified, but not before the dispute was dragged into the High Court and finally settled in



Zoo guests

the Court of Appeal. The proposals and court ruling were incorporated in a new Royal Charter in 1963.

Ivor Montagu, one of the ZSL's Vice-Presidents and a regular visitor to the Zoo for nearly seventy years, believes that the 1957 proposals were absolutely essential for the Zoo's survival. In his view the strength of the institution lies in its Charter but the composition of the Council has sometimes made it vulnerable. He says it was set up in a way which grew alien to the world around it: the Zoo in its early days was very much a preserve of the gentry; a walk in the Zoological Gardens on Sundays by the Fellows and their friends was very much 'the thing to do'. The First World War altered the social climate and this, he says, forced the Fellows to relinquish some of their privileges. By pushing the 1957 proposals through, he argues, Lord Zuckerman "rescued the Zoo and brought it into the modern world." For many years London Zoo had indulged in the policy of 'stamp-collecting'—it had to have at least one specimen of each species. But with greater concern about conservation and the alteration in public attitudes that policy changed. The emphasis is now placed on maintaining fewer varieties of animals and these in viable breeding colonies.

Zoo activities

The ZSL is an institution which maintains the national collection of animals. It owns one of the most comprehensive zoological libraries in the world and is responsible for an extensive educational scheme which catered for 53,000 pupils in 1976. The Society publishes more scientific journals than even the Royal Society and

in addition it runs two research institutes of world renown. These activities, undertaken in addition to the normal display of animals, are financed in the main by gate proceeds. The ZSL's 1976 Annual Report states that with a turnover of some £2.5 million and a combined attendance of close on 2½ million at Regent's Park and Whipsnade, the Society had a year-on surplus of £179,000. But this sum, unfortunately, will have to be used to pay off the deficits accumulated in 1974 and 1975 when the effects of inflation and recession were sorely felt.

The additional activities undertaken by the ZSL are an obvious strain on the Society's revenue. Both in Britain and abroad many other institutions with far fewer responsibilities are in receipt of regular government subventions. London Zoo is not so endowed. The obvious question is whether it should be. Lord Zuckerman is circumspect with his answer. All he says is that the ZSL is as worthy of support as many of the institutions in regular receipt of public funds.

An intense debate is being conducted at the moment by the staff within zoos to decide their real purpose—the educational function of zoos, their role in conservation of endangered animal species, the nature of scientific research to be carried out and so on. The 1976 Report acknowledges part of that discussion. With its facilities the Society is well placed to become "an international centre for the scientific study of the problems of conservation." The Council proposes to combine the two research institutes, the animal hospital and pathology laboratory into an integrated 'Institute of Zoology'. The scientific expertise and support of the

new institute in reproductive physiology, biochemistry, immunology, radiology, parasitology and genetics "can put real knowledge behind attempts to breed stocks of rare animals". This is to become one of the institute's main projects.

Finance for conservation research, however, is very difficult to find. Many zoologists feel that at least some of the money raised for wildlife preservation generally could be used more profitably for the scientific research needed to back up the conservation programme. One organisation concerned with conservation is the World Wildlife Fund. Its Vice-President and Founder Chairman, Sir Peter Scott, feels that conservation is important in the context of zoos. He says zoos are especially important in two respects: for the captive breeding of species threatened by extinction, and for education. Zoos have a great responsibility to educate the young, he argues, and the ZSL has done well in both areas; although the Fund did not at the moment finance any conservation research at London Zoo, if a good case was made they "would consider supporting it".

The ZSL's total income in 1929 was about £12,000. Some 16% of this was spent on scientific activities. Today the proportion of income spent on research, publications and education stands at the respectable level of 18%.

Hedley's task

Ronald Hedley, the man succeeding Lord Zuckermann as Secretary of the ZSL, is a biologist by training with a background in chemistry, zoology, geology and physiology. His experience at the Natural History Museum should equip him well to deal with the public and with the scientific and educational functions of the Zoo; last week he said he was looking forward to the prospect of working with new colleagues there. But he is not approaching his new job with any preconceived ideas. Believing in the need to sample opinion, of which there will be no shortage at the Zoo, he intends to have a period of consultation first.

The Council of the ZSL has the ultimate responsibility for ensuring that the Society is successful. Dr Hedley as a member of that Council now has the two-fold task of involvement in the running of the Zoo and the Natural History Museum, which is an essentially complementary institution. He is conscious that the goodwill and support of their staff is essential for their continuing success, and takes up his post when the Zoological Society of London is probably more confident about its future than at any time in the past fifty years. □

As of this week Lord Zuckerman will preside over all the Zoo's affairs and chair any of its meetings which he attends. Previously, as the Zoo's Secretary, he had a more demanding task, which is one obvious reason why he might retire from the post. But, though now 73, he will still be around to offer advice, just as he has done in another capacity, that of a government science adviser. This link has lasted for the best part of 40 years, and culminated in promotion to Chief Scientific Adviser from 1964 to 1971. He is still to be found in and around the corridors of power.

Interviewed last week, Lord Zuckerman responded firmly to the suggestion that with the explosion of information no single individual could possibly advise over the whole spectrum of science. What was wanted of a scientific adviser, he argued, was an understanding of the "political, national and international" implications of technical decisions taken by governments; the immense political implications of President Carter's decision not to proceed with the breeder reactor, for example. That a political sense was required was particularly true

when all government departments already had specialist technical committees to consider long term planning decisions. Zuckerman says that "the idea of appointing a scientist just because he is a Nobel Laureate who has unravelled the structure of some particularly heavy protein molecule is nonsense. The idea of having the best mathematician is nonsense".

It needs no mathematician to calculate that Britain's Research Councils are now having a comparatively lean time. Zuckerman acknowledges that in the 1950s and 1960s he was one of the individuals who convinced people that science should be supported in a much bigger way and that now people have the right to enquire after the results of that policy. With the golden age of expansion over, some scientists will draw little comfort from his belief that they will just have to accept the cuts as a fact of life in a country as "hard-up" as Britain. But he believes the current period of belt-tightening will not restrict "real creative thought", and that there will be no long-term repercussions for scientific output.

Fiftieth anniversary of the Dunn School of Pathology



- 1 The Sir William Dunn School of Pathology
- 2 Dr Jean Medawar, Dr Anne Beloff-Chain, Lady Florey, Sir Peter Medawar, Sir Ernst Chain, Professor A. D. Gardner
- 3 Dr George Mackaness
- 4 Professor Dorothy Hodgkin, Professor D. Whitteridge
- 5 Professor Henry Harris



WHEN the trustees of Sir William Dunn's bequest offered to support the founding of a new pathology building in Oxford they stipulated that it should be fronted by flower beds. On 20 April, the occasion of the 50th anniversary of the opening of the Sir William Dunn School of Pathology, only the reminiscences of the plentiful past members of the School were more colourful and varied than the flowers outside. The celebrations were launched, in a style unequalled by later speakers, by Emeritus Professor A. D. Gardner. An Oxford graduate of 1907, Gardner recalled how Sir William Dunn's money, earmarked primarily to aid the destitute and the orphaned, came to be used for funding medical research. The School's first director was Professor Georges Dreyer whose strength, said Gardner, lay in his thorough statistical approach to research which was centred on the development of tuberculosis vaccines. His weakness lay in his belief that he was infallible—a vice, Gardner emphasised, that was unique neither to the man nor the era.

After Dreyer, in 1935, came Professor Howard Florey who was remembered in his pre-war days by Sir Peter Medawar for being exhilaratingly energetic and exasperatingly anti-intellectual. The war-time saga of the isolation and first therapeutic trials of penicillin by Florey's team was narrated by Dr Norman Heatley who emphasised the contributions of novel hardware (bed-pans, biscuit tins and Russian butter tins) and novel,

hard-wearing female technicians. In 1952 when Henry Harris arrived from Sydney, Florey's direct interests in antibiotics were being wound up in J. L. Gowans' DPhil thesis. Thereafter antibiotic research was carried out under E. P. Abraham who told the story of cephalosporins from their bizarre origins in a Sardinian sewage outlet to the present day.

Florey's interests, ruled as ever according to Professor Harris by the illnesses with which he or members of his family were afflicted, turned to atherosclerosis, gastric secretion and inflammation. A three-pronged attack on inflammation was mounted by Florey through George Mackaness, Gowans and Harris who were to investigate the roles of the macrophage, the lymphocyte and the polymorph, respectively. Only Harris gave up. His decision that the polymorph had been good for a DPhil but would not do for a career brought out the fiercest in Florey. Harris was threatened with deportation to Australia, and communication between the two became confined to letters. Fortunately Harris's colleagues, particularly Abraham, rallied round and persuaded Florey to back down. Henry Harris survived to take over Florey's chair in 1963. We look forward to the 75th anniversary of the Sir William Dunn School of Pathology when tales of Henry Harris can best be told, perhaps under the shade of the ginkgo tree ceremonially planted by his wife on this occasion.

Peter Newmark

USA

The costs of high technology

A report last week on X-ray scanners has raised questions about expensive technology in the US health care system. Colin Norman reports from Washington

IN mid-1973, a sophisticated and very expensive piece of medical equipment was introduced into the United States from Britain. Since then it has found its way into hundreds of hospitals and doctors' offices, and the swiftness with which the new technology has been adopted has generated a major controversy. This has helped shed light on the reasons why hospital costs have risen eight times faster than the cost of living during the past 25 years, and why many people despair that the American health care system can ever be brought under sensible cost controls.

The equipment, known as a computed tomographic (CT) scanner, is a nifty machine which its promoters claim will revolutionise the diagnosis of many medical disorders. Developed by EMI Ltd, the CT scanner is essentially a combination of an X-ray machine and a computer, which provides cross-sectional images of internal body organs. Early models were limited to use on the head, but later and more sophisticated versions can provide images of the internal organs of any part of the body. The scanner's advantage over conventional X-ray machines lies in its ability to produce high quality images of soft tissue structures from numerous angles.

Few people would quarrel with the argument that CT scanners represent a major advance in medical X-ray technology, or that they may be useful diagnostic tools in many instances. The controversy centres on the fact that the technology has been so widely and swiftly adopted in the United States that massive investments have been made before the clinical usefulness of the machines has been well documented.

Although less than four years have elapsed since the first CT scanners were introduced into the United States, an estimated 760 machines have now been installed or are on order. Since each machine has a price tag of between \$300,000 and \$700,000 and costs about \$300,000 a year to operate, hundreds of millions of dollars are being sunk into the technology by American hospitals. A particularly striking indication of the overwhelming demand for the machines is the fact that some 40 orders were placed for EMI's new whole body scanner before the machine

was even commercially available. By contrast, in Britain, where CT scanners were originally developed, only a handful of machines are in operation and fewer than 40 have been ordered.

This overwhelming and seemingly unreserved adoption of CT scanners in the United States led the Health Research Group, an organisation linked with the consumer advocate Ralph Nader, to charge last year that "the purchase of these machines is a particularly striking example of the introduction of extremely expensive medical technology into the health care system before there is adequate evidence that patient benefits or dollar savings exceed costs". The group demanded that a national moratorium be placed on the purchase of CT scanners until the benefits had been better evaluated. And last week, the Institute of Medicine, part of the National Academy of Sciences, said in a carefully worded report that "the long-term effects of CT scanning on medical care and its costs are not yet discernible".

The Institute's report, perhaps the most influential study yet published on the use of CT scanners, is noteworthy as much for what it says about the general problems of controlling hospital costs as for its recommendations about this particular technology.

The report is quick to acknowledge that CT scanners may have many important uses, and it lists some of the areas in which the machines may offer advantages over conventional diagnostic tools. Nevertheless, the Institute notes that conventional techniques may be just as effective and much cheaper in many cases, and "considers it essential . . . that controls be placed on the location and use of and charges for CT services". It also recommends that medical insurance companies should not pay for some types of CT scan until there is better evidence of their efficacy.

The question arises as to why the American hospital system has rushed to embrace CT scanners so wholeheartedly. The answer lies in the organisation and funding of the health care system, and the swift adoption of CT scanners epitomises many of the problems inherent in that system.

To begin with, there is little political incentive to hold down the costs of medical care in the United States. The middle classes, where political power largely resides, are generally well protected by private health insurance, and the medical costs incurred by the poor and the aged are met by the federal government. As the Institute puts it in

its report, "The growing share of personal health care expenditures covered by third-party reimbursement has reduced the incentives to control use (of expensive technologies), because the physicians decision to use services is separated from the patient's immediate expenses". According to officials in the Department of Health, Education and Welfare, more than 90% of the \$55,700 million hospital costs incurred last year were met by private insurance payments or federal Medicare and Medicaid schemes.

The report cites other reasons for the soaring costs, such as "competition among hospitals for medical staff, prestige, or revenues", and it notes that there has been a recent general increase in the use of diagnostic tests stemming partly from fear of malpractice suits, which "encourages defensive medical practice". All of these factors, the Institute states, suggest that "the present organisation of medical care and methods of financing and regulating that care in the United States have encouraged investment in beds and equipment beyond a socially efficient level".

Some states are trying to get a grip on soaring hospital costs through so-called certificate of need laws which require institutions to obtain approval from planning authorities before purchasing expensive equipment. But these laws, for a variety of reasons, have failed to curb the proliferation of CT scanners. The Institute report notes, for example, that some 15% of the CT scanners installed so far are located in the offices of private physicians, which are exempt from the certificate of need requirements, and "some physicians appear to have deliberately circumvented the intent" of the laws. In one well-publicised case, for example, after a hospital was denied a CT scanner, physicians at the hospital formed a partnership, bought a machine and installed it in an office across the street.

A final incentive for hospitals to purchase CT scanners seems to be their potential for generating revenues. The Institutes report states, for example, that the average charge for a scan is about \$240, and an independent economic survey of CT scanners published last year indicated that the average machine is used for about 3,000 scans per year. In other words, about \$700,000 is generated by each machine per year, more than enough to offset the purchase price and operating costs over a very short period.

Given that profit potential, there is some incentive for hospitals to use the machine in cases where less costly, conventional diagnostic equipment would

USA

● Total spending on research and development in the United States is expected to reach nearly \$41,000 million this year, according to figures compiled by the National Science Foundation. If the estimates are correct, support for research and development would stay ahead of inflation for the second year running, though the combined two-year increase would not be sufficient to wipe out inflationary losses incurred during the early 1970s.

According to the NSF study, the federal government is expected to spend about \$21,800 million in 1977, a 10% increase over 1976, while industry is expected to spend about \$17,500 million and about \$1,500 million will come from universities, non-profit institutions and so on. In terms of constant dollars, the total is expected to be about 6% below the peak spending years of the late 1960s. Measured as a proportion of the gross national product, spending on research and development has declined steadily since 1964.

NSF has also published an estimate of the number of scientists and engineers employed in 1976, which suggests that after four years of growth, the scientific labour force now numbers 542,000. The total is about 20,000 below the 1969 level, however. Some 40,000 science and engineering jobs were lost during the 1969-1973 cutbacks, NSF reckons.

● While the nuclear industry in the United States was licking its wounds last month, following President Carter's decision to defer commercial reprocessing and to downgrade the breeder reactor programme, a report on nuclear power plant security drew some fresh blood. Published by the General Accounting Office (GAO), an investigatory agency of the Congress, the report concludes that "security systems at perhaps all power plants would not be able to

withstand sabotage attempts by threats that are now considered minimum by (the Nuclear Regulatory Commission)".

Based on inspections of security systems at six power plants, carried out by GAO investigators, the report is written in uncharacteristically blunt language, laying blame at the doors of both the nuclear industry and NRC. It states that though the quality of security forces seems to vary greatly from plant to plant, some deficiencies were encountered on each inspection. The faults included lack of training of security guards (as little as four hours at one plant), high turnover of security personnel (up to 48% a year), and poor security equipment. In particular, the report cites the following two horror stories:

We accompanied an NRC inspector to one power plant at night. The inspector asked the guard manning the guardhouse to aim a closed circuit television camera on a particular spot. The guard tried but was unable to work the system. The inspector opened a door which rang an alarm in the guard house. After waiting several minutes, the guardhouse was called to find out why no one responded to the alarm. A guard in the guardhouse answered that all of the available guards were too busy.

At yet another site, we asked a guard about the locations of certain critical systems of the plant, including the control room. He told us that the guard force knew nothing about the location of these systems because the guards were not allowed inside the power plant.

Although the NRC has recently published a set of new regulations designed to increase the effectiveness of security systems at nuclear plants, the regulations are not due to come into effect until August 1978. The GAO report says that although the new regulations are "on the right track", NRC should take steps immediately to ensure that operating plants are made more secure.

Colin Norman

serve just as well. To try to limit such potential abuses, the Institute recommends that medical insurance companies should establish an advisory committee to develop criteria for paying CT scan charges. It also recommends that requests for use of CT scans should be reviewed by a physician with responsibility to control access to the machine, who would determine whether the scan is appropriate.

The central question here is not whether CT scanners are useful, but what level of investment can be justified and how the machines should be allocated and managed. Those are questions that the American health

care system finds difficult to address, given its fragmented nature, the competition between hospitals for revenues and prestige and the third-party payment system. As Dr Harvey Fineberg, Director of the Graduate Program in Health Policy and Management at Harvard School of Public Health put it last week, "The fundamental assumption that we in this country have not yet made is that the pot of resources available for health care is limited . . . if you ask whether CT scanners are useful, the answer would be 'yes', but if you frame the question 'would they be a sound investment', you may come up with a different answer". □

WEST GERMANY

Energy plan published

Werner Gries reports from Bonn on the energy research programme recently approved by the West German government

THE main emphasis of the German federal government's DM 6,200 million energy research programme for 1977 to 1980, published at the end of last month, is on nuclear research, but the rational use of energy, progress in coal technology and the development of new energy sources are also prominent. Details of the programme, which marks the first time the state has intervened to bring together systematically both nuclear research and work on other energy resources, are as follows:

Nuclear energy will take some 75% of the state funds. Fast breeder and high temperature reactors remain the first priorities in the country's nuclear power programme, but research and development work is to continue in the fields of uranium enrichment, reprocessing and the disposal of radioactive waste. A state subsidy of about DM 1,200 million will be needed just to complete the prototype power stations now under construction to house the breeder and high temperature reactors. Reactor safety is a key aspect of the programme; R&D on nuclear-powered ships will be continued only on a fairly modest scale, receiving about DM20 million annually.

Rational energy use: Around DM100 million annually will go into research into technologies for rational energy use. Attention will be given to techniques of remote heat supply, reverse cycle heating systems, heat recovery processes and the use of waste heat from power stations.

Coal technologies: R&D in coal technology, one of the most important aspects of the non-nuclear side of the programme, will receive increasing amounts of funding, amounting to an average of DM140 million annually. Coal gasification, improvements to coal-fired power stations to reduce pollution, coal liquefaction and improved mining techniques will all receive support.

New energy sources: The government will support the development of new sources of energy to the tune of DM130 million a year on average, with the focus on nuclear fusion (about DM90 million annually) and on solar energy. In the government's view, solar energy is of interest in West Germany only as a means of heating build-

ings and providing hot water, and research is being concentrated in this area. Further work on nuclear fusion is dependent upon the site of JET; the Research and Technology Minister, Hans Matthöfer, expressed his hope for an early decision on this.

It is Herr Matthöfer's ministry which has much responsibility for R&D in West Germany. Its budget in 1977 of DM4 million million compares with total research expenditure in 1977 estimated at DM26 million million. Federal research programmes are designed for a period of about four years, classified by sectors and provided with a corresponding budget during the planning stage. The energy programme is one of ten programmes for

sectors that also include data processing, marine research and technology, raw materials, electronic components, space and aviation, and so on.

The outlines of the new energy programme were determined by the federal government following a cabinet resolution. The programme is not immediately binding for industry but outlines the frame of what the government considers essential. The country's primary energy consumption in 1975 broke down as follows:

Mineral oil	52.1%
Hard coal	19.1%
Soft coal	9.9%
Natural gas	14.0%
Nuclear energy	2.0%
Others	2.9%

The structure of energy consumption differs from the United States and other large countries. A breakdown of energy consumption by sectors in 1975 shows that industry consumed 35.9%, transport 19.7% and domestic 44.4%. In the government's opinion the largest savings have to be achieved in the domestic sector. Industrial energy consumption, related to national product, has actually stagnated for about ten years.

An output capacity of 33,000 MW from nuclear power stations in Germany was anticipated for 1985, but a maximum of only 25,000 MW can be expected because of recent breakdowns in power stations and the postponement of various building projects. □

SWEDEN

● A report recently published by Sweden's National Environment Protection Board shows that the business of making polluted lakes clean is slow and difficult. Sweden has 100,000 lakes covering 9% of the country's surface, and of these about 15% in the southern third of the land are being overloaded with phosphates and nitrogen. In general, the polluted lakes are victims of acid rain, sewage, industrial effluents and agricultural run-off. Recovery programmes are having limited success.

In the long term, acidity will be reduced only if fuels can be burned without releasing much sulphur into the air; but as a stop-gap measure, lime has been spread onto the lakes themselves in some severely-affected areas. Prospects for large-scale liming seem poor, however, not only because of the cost but also because the traces of toxic metals found in most commercial lime would become significant in the quantities involved.

Another recovery tactic has been to cut off flows of effluents into a lake and the streams feeding into it. This has only improved water conditions slightly, as the lakes' internal circulation of nutrients has not been affected enough to cut down the production of plankton to acceptable levels. Reducing pollution by treating sewage to remove the nutrients from it is not always effective either: although it caused plankton levels to go down in one case, the gains were partly offset because the farmers in the area increased their use of nitrogen fertiliser.

In some cases, causes and effects in the pollution cycle are impossible to determine. The fact that the burbot in Lake Vättern has decreased dramatically in the past ten years while its main prey species has increased could be due to the burbot's

high PCB content or simply to overfishing. But once the balance between the two species has been upset it may never be re-established, and the reason for the change may never be known. Pike in Lake Vänern are black-listed because they contain mercury from industry at the northern end of the lake. The fact that they have been fished less has meant that their average size has grown, which has resulted in increased cannibalism and a slower growth rate in their population. In these circumstances it is extremely difficult to say exactly what the direct effects of mercury pollution on pike may be.

With these uncertainties, recovery programmes are hard to design. The report gives the impression that the problem would be easier to tackle by prevention rather than cure.

● The Nuclear Fuel Supply Company has reached an agreement with the French firm Cogema under which spent nuclear fuel from the Barsebäck 2 and Ringhals 3 reactors will be reprocessed at the French facility in La Hague. Cogema will accept spent fuel produced until the end of 1979, and will return it after reprocessing to Sweden for final storage. Because of delays in the Ringhals 3 programme, it seems likely that the spent fuel sent will in fact come only from Barsebäck 2.

Under the law specifying the conditions to be met by reactor operators before the government will give permission for the building and operation of nuclear power plants, the owners of Barsebäck 2 must organise secure reprocessing of spent fuel. In their opinion the present agreement fulfills their obligation. Sweden's Energy Minister, Olof Johansson, who is trying to stop Sweden's march into the nuclear society, is refusing to admit

defeat. He says that the government can decide whether the owners have met the conditions only after it has examined the contract. The government's pronouncement is expected to be made by 1 October.

● According to a recent statement by the Minister of Education, Mr Jan-Erik Wickström, Sweden may be facing an acute shortage of people educated in technology and the natural sciences. The 1970s, said Mr Wickström, have seen difficulties in recruiting students to technical and natural science lines at both secondary and tertiary levels. He blamed this partly on the state of the labour market, which was unfavourable for natural science graduates at the beginning of the 1970s, partly on the evident inability of primary schools to exploit students' interest in natural sciences, and partly on the unpopularity of technicians and scientists in a society affected by the green wave.

A look at the official statistics hardly clarifies the picture. It is true that the number of students reading natural sciences and mathematics at tertiary level fell by 25% between 1971 and 1976. But that should be seen against a background of a general drop in interest in tertiary education: the number of students reading humanities and social sciences, for example, fell by 20% during the same period. Yet these years saw the number of students at institutes of technology increase by 9%. A similar pattern holds for doctoral degrees taken over the same period: down by 16% for mathematics and natural sciences, down by 17% for humanities and social sciences, but up by 5% for technology.

Wendy Barnaby

IN BRIEF

Nuclear study

In a declaration issued at the end of the two-day summit meeting in London at the weekend, the leaders of the world's seven leading non-communist industrial nations committed themselves to increasing nuclear energy to "help meet the world's energy requirements" and announced the launching of an "urgent study" (a "preliminary analysis, to be completed within two months", according to an annex) on how to do this while reducing the risks of nuclear proliferation.

The announcement, one of the main features of the final declaration, was carefully worded to smooth over the differences that have emerged within the western alliance since President Carter began acting on his worries about international trade in nuclear technology for ostensibly civilian purposes.

The proposed committee is expected

to consist of experts from the various countries and the study, on uranium supplies and aspects of the nuclear fuel cycle, will decide the terms of reference to be followed for a broader study lasting perhaps a year. Though it will be difficult for any such committee to agree conclusively, particularly in so short a time on what is now so political a matter, the key point is itself political, namely that the study represents a review of nuclear policy at the highest possible political level.

Geos update

Initial testing of experiments on board ESA's research satellite Geos has now been completed; all were working well early this week. Final operating conditions were due to be achieved by mid-week when the long radial booms would be fully deployed to 20 m. The booms had been at 2.5 m, giving a rather high spin rate for the particle

experiments. The slow drift of the apogee to longitude 37.7°E was halted on 8 May and Geos is now in an orbit with apogee conjugated to Northern Scandinavia, but a limited programme of correlated measurements will be possible.

The satellite is in view on every second 12-hour orbit for nearly 8 hours per day and the viewing time is well matched to the period during which the altitude is greater than the 3.5 Earth radii where all experiments can operate. The viewing time could be doubled by collecting data from the other orbit using NASA ground stations in Alaska, but the apogee remains at the same local time. It will be next February before the local time of the apogee moves from early evening through the day to probe the night-side region of the magnetosphere.

NASA's investigation into the failure of the launch vehicle continues. A detailed report is expected this week.

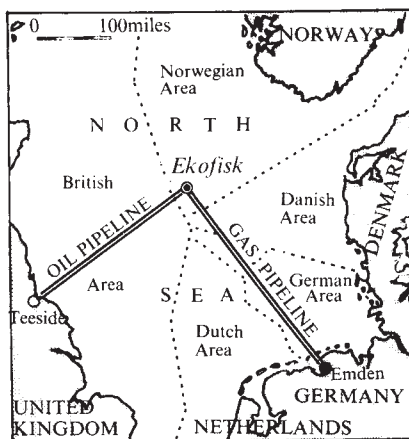
ON 30 April the blown-out Ekofisk Bravo oil well in the middle of the North Sea was finally capped, having spewed 12,000 tons of oil into the sea in the previous nine days. Three days later the oil, constantly tracked by satellite, had virtually disappeared. Such a startlingly anti-climatic ending to a drama billed as an ecological disaster must leave a few people somewhat puzzled as to where it all went.

A good deal of it of course went upwards. The UK Department of the Environment in *Accidental Oil Pollution of the Sea* (Pollution Paper no. 8, 1976) estimates that up to 40% of the relatively light North Sea oil may evaporate in one day; oil from the Ekofisk field has an even lower specific gravity than the North Sea average. The volatile fraction fortunately includes the most toxic components of the oil, the aromatic hydrocarbons such as benzene and toluene.

What was left after evaporation and solution formed that water-in-oil emulsion known as 'chocolate mousse' whose persistence normally causes great despondency. But in this case, probably again because of the lightness of the oil, the emulsion was an unstable one, quite easily dispersed by the rough seas which had conveniently developed soon after the well was capped. The silver sheen slick of dispersed oil droplets covered a huge area for a short time, but by the middle of last week was reduced to a few square miles.

Once the oil had been broken down into small droplets the paraffins could be digested by bacteria. Some particles are ingested by zooplankton which die and fall to the seabed, making the oil available to the more abundant bottom-living bacteria. The

Where goes the oil?



BACKGROUNDER

enormous amount of extra carbon available will cause short-term changes in the community but probably no long-term disruption. Though normally one would expect some tar to be formed from the heavier residues there may be virtually none from the Ekofisk oil.

Nature was responsible for 90% of the oil dispersal; the rest was managed by the Norwegians, who contained the oil with booms and

managed to skim 1,000 tons off the surface. A better figure probably could not have been hoped for: removal of the oil is certainly the ideal solution to the pollution problem but present equipment, being ineffective when waves are more than 7-foot high, is useless for most of the time in the very rough North Sea. The clean-up ships were fortunate in this case to get two days of near calm to pump most of the thicker patches of oil off the sea.

By yet another happy chance the blow-out happened at such a time as to cause minimal damage to fisheries and birds. The Ekofisk field is in the middle of Norway's valuable mackerel fishing ground; mackerel do not swim near the surface but their eggs, which are laid during May, float near the surface and would certainly have been damaged if the blow-out had happened three weeks later. The Norwegians are reluctant to use chemical dispersants on oil because of their toxic effects on marine life. In this case they were proved right by the rapidity with which the oil disappeared but there would have been strong pressure to use chemicals if the accident had happened in July, when large numbers of auks are swimming across the North Sea from Scotland to Norway.

Thus, apart from a few unanswered questions, for example concerning the long-term effects of hydrocarbons in food chains, there is cause for great relief. Things may not work out quite so well next time.

CEGB change

The retirement of Sir Arthur Hawkins as chairman of the UK Central Electricity Generating Board (CEGB) at the end of last week has raised the prospect of an early resolution of the problem which the proposed Drax B coal-fired power plant poses. The issue is not whether it should be ordered, but when. That depends in the first instance on electricity demand, but other factors have intervened.

The plant was planned for 1979. The turbine and boilermaking industries face a lean time with demand down and want the order brought forward; the government's 'think tank' has recommended a rationalisation of the industry and this is still under discussion. The CEGB has said the order can be placed now, but wants compensation from the government. What is in dispute is the amount of that compensation, and who should pay it. The hope is that with the departure of a key

figure in the growing controversy, some new formula can be found to satisfy Glyn England, the new CEGB chairman. With the visit of Mr Benn, the Energy Secretary, to Moscow this week for talks with Mr Kirillin, his Soviet counterpart, no announcement is expected before next week.

On the move

Dr Frederick Horner is to take over from Dr J. A. Saxton as Director of the Appleton Laboratory near London. Horner joined the Radio Division of the National Physical Laboratory in 1941, which has since evolved into the Appleton Laboratory. In 1969 he became Deputy Director. Most of his research has been concerned with finding radio direction and studying radio emissions from thunderstorms. Since 1950 he has been involved in international radio affairs and has held office in international organisations. Saxton retires on 1 July.

The UK Department of Health will soon be losing its Chief Scientist, Sir Douglas Black, following his election as President of the Royal College of Physicians. No replacement has been announced.

WHO message

The World Health Organisation (WHO) should strive to achieve a universal level of health conducive to high social and economic productivity by the year 2000. This was the message of Dr H. Mahler, Director-General of the WHO, when he presented his report for 1976 to the World Health Assembly in Geneva last week. He said that the WHO should through its country health programming promote development in primary health care and appropriate health technology, even though in so doing it might meet with opposition from the medical and various allied professions and industries.

IN Britain, most western European countries and the United States, the so-called 'organic movement' is growing. Its members believe that 'natural' manures, derived from animal excreta or composted vegetable matter, are better than factory-made chemical fertilisers. The movement embraces those with widely different views. The more extreme believe that chemical fertilisers poison the land and those who eat its products. The more moderate consider that organic manures improve the structure and fertility of the soil better than do inorganic fertilisers, and produce crops which differ nutritionally in subtle ways from those grown with liberal dressings of artificials. They claim that pests and diseases are less rampant when organic techniques are used, though they admit that the highest yields are often obtained with balanced dressings of inorganic fertilisers and the liberal use of chemical pesticides. In general they claim quality rather than quantity for organic produce.

As an entomologist on the staff of Rothamsted Experimental Station, on whose fields the classical experiments demonstrating the value of chemical manures were performed, I was at first very sceptical. However, when, over twenty years ago, I visited the experiments of the Soil Association at Haughley in Suffolk, and examined the long-term experiments set out and organised by Lady Eve Balfour, I decided that this work needed to be looked at critically and scientifically. I did not have an experience like St Paul on the road to Damascus, with an instant conversion, and I found

some of the explanations of organic farmers and gardeners hard to accept, but I did think that this unique long term study, where comparable fields were receiving different fertiliser and pesticide treatment, ought to be thoroughly investigated.

I attempted to interest bodies like the Agricultural Research Council,

Muck is magic



KENNETH MELLANBY

who were at that time comparatively generously supplied with funds to support research. I could not persuade them to part with a penny to investigate the work at Haughley, or to study organic methods generally. Time and time again my advances were rebuffed. Even when I got together a Research Advisory Committee, which included hard-headed

scientists from the chemical industry, no government money from agricultural research funds was forthcoming. In fairness to the ARC and my former colleagues I must admit that they believed that the subject had been thoroughly studied by established methods, and it was not unreasonable to assume that little more was to be discovered. Furthermore, some of the extreme organic fanatics talked such evident nonsense, and their more rational colleagues suffered by association.

Today things seem to be different. I was delighted to see a headline in the *Times* recently which read "Ideas of organic farming being taken more seriously", and to discover this referred to a lecture by a member of the staff of the ARC. Not long before I heard my former colleague Dr G. W. Cooke, recently the distinguished head of the Chemistry Department at Rothamsted and now also with the ARC, admit that modern work had shown that organic manures were now being found to increase some crop yields much more than would have been expected from an analysis of the inorganic nutrients present.

I hope that this will mean that organic methods will now be really critically studied, as I proposed so long ago. If the ARC sees the light in this way, they may wish to follow the example of the now-defunct Council for Scientific Policy. This body prefaced its last report with an extract from the General Confession in the Book of Common Prayer. I suggest that the ARC choose as its text Luke 15, v. 7.

correspondence

'The Selfish Gene'

SIR,—Lewontin's review of Dawkins' book *The Selfish Gene* (17 March, page 283) is a disgrace. It fails to meet any of the standards of informative value, objectivity and fairness to the views of others that are part of the code of science. That it is a very biased review is reasonably evident to anyone reading it; unfortunately a reader unacquainted with the controversy which is its background may well be left with the impression that, even setting aside obvious unpleasantness in the review, the book itself probably is unsound and not worth reading. This is a great pity since in fact the book is not only the best existing outsider's introduction to a new paradigm and a new field of knowledge but, in its overview of the situation and in many original details, is itself a significant contribution to this field.

For its intellectual worth, and seemingly in motivation as well, Lewontin's outline of the book is on par with Bishop Wilberforce's notorious attack on Darwin and Huxley at the British Association meeting of 1860. It would be easy to defend the book and reply to the review point by point, but this would be an unpleasant, muck-raking task, and it would be difficult to avoid matching rhetoric against rhetoric and so bringing debate to the plane where Lewontin wants it to be. Against such a review I feel justified in following Huxley to give a very subjective assessment of my own but wish to avoid trying to match the wit of the historic reply to the bishop.

Whether the literature of socio-biology reflects an ignorance of the difference between "properties of sets

and properties of their members" or "confusion between materialism and reductionism" are matters about which I feel little concerned. On the other hand I feel a warmth very far from indifference when I encounter in most of the literature in question signs of a spirit which I share and which I have always assumed is the same as that which motivates scientific enquiry in all its branches. I can most simply express this spirit by calling it a desire to understand and communicate the nature of the world. I find it present in full measure in *The Selfish Gene* and totally absent in Lewontin's review.

Yours faithfully,

W. D. HAMILTON

Imperial College Field Station,
Ascot, UK

Efficient fish

SIR,—"... fish are not efficient users of food materials other than animal protein and fat", according to Dr Purdom and Mr Preston (31 March, page 396). This is incorrect for some species. At Reading we routinely obtain feed conversion ratios of less than 10 kilograms of fresh vegetation to 1 kilogram of fish when feeding ryegrass (*Lolium* spp.) to grass carp (*Ctenopharyngodon idella* Val.). This represents a protein conversion efficiency of 40%. Equally efficient conversion of terrestrial vegetation is recorded in the literature for *Tilapia melanopleura* Dumeril and *Tilapia zillii* Gervais in pond culture. The figure of 100 to 1 given by Purdom and Preston is applicable to a diet of aquatic vegetation which consists largely of water. Moreover, we have evidence which suggests that leafy ter-

restrial vegetation provides all the major nutrients required to support rapid growth rates in grass carp.

If it is assumed that ryegrass for artificial drying costs £6 per tonne fresh weight, the feed cost is 6p per kilogram of fish produced. Feed costs could be reduced by using agricultural by-products and food-processing wastes. It should therefore be technically possible to farm fish intensively and at low feed costs. Farmed herbivorous fish are rated highly as food throughout much of Asia, Africa and Eastern Europe.

Yours faithfully,

J. C. J. DOMANIEWSKI

Department of Agriculture and
Horticulture,
University of Reading, UK

Value of textbooks

SIR,—During recent discussions on the structure of the living cell at meetings of the Physiological Society, the Anatomical Society, the Biochemical Society, and the Royal Microscopical Society, several participants have made the remark "Surely, you don't believe what is written in text books". This viewpoint seemed to be held by several other participants at the meetings. This raises such fundamental questions of scientific truth, that we wonder if any of those who support this view might be prepared to justify it in print. We should say that we were grossly disturbed by such a suggestion and thoroughly disagree with it.

Yours faithfully,

H. HILLMAN

P. SARTORY

Department of Human Biology and
Health,
University of Surrey, Guildford, UK



Competition 13

(From an idea by Martin Rees). In the next few years we discover—never mind how—that there is intelligent life nearby in the universe, but that we are more 'advanced' than they. A prize of £10 is offered for the best short message to their equivalent of

Isaac Newton, or any other character from our history, telling him what scientists are up to these days. Send entries by 15 June to the Editor, and mark envelope Competition 13.

Competition 12 asked for fatuous, actual or fictitious professorships. Honourable mention to Drs Haslam and Haslam for the fictitious Chair of Wind Engineering in the Medical Faculty (to go with a new journal entitled Wind Engineering) and the actual Chair of Artificial Intelligence.

But the prize to Professor V. Trimble (Irvine) for the following:

It would be difficult to top Cambridge's own (real) Plumian Chair of Astronomy and Experimental Philosophy, whose incumbent, we presume, is required to go about kicking stones and looking for trees in quads where there is no one about to observe them. Recent interest has, however, also been attracted by Brown University's (fictitious) chair of Psychoceramics, whose present holder, Professor Carbury, is, of course, an expert on cracked pots. My own university's English department has contemplated a chair in Historical Metallurgy to be reserved for someone learned in the study of Old Saws.

news and views

Cosmic gamma-ray bursts

from K. J. Orford

SINCE their first reported discovery 4 years ago, cosmic gamma-ray bursts have been observed at a rate of roughly one per month. These events have until now been observable only from satellites, or balloons, since the weak gamma-ray energy flux cannot penetrate the atmosphere or compete with atmospheric radiation. Their theoretical interpretation has been hindered by a lack of sufficient knowledge of their celestial positions, their distribution in size, and a lack of any correlated phenomena which could narrow the search for sources. As a consequence the number of theoretical models has exceeded, if not the number of bursts (about 50), then at least the number of well established parameters. Some recent results reported in *Nature* may help.

Manchanda and Ramsden (*Nature* **266**, 425; 1977) point out that about 90% of the bursts have sizes, or energy fluxes, within a relatively narrow range. If the assumption is made that they all originate in explosive events of similar energy output then a burst's size reveals its distance, at least relatively. The distribution in distance so obtained bears a remarkable resemblance to the distribution of supernova remnants. Supernovae are thought to be sources of both neutron stars and black holes. The disruption of a neutron star by a black hole would provide a neat explanation not only of the size distribution, but also of the gamma-ray burst's energy spectrum. Two consequences of this theory are that the bursts should have a particular distribution in celestial position and that very small and very large bursts should be relatively infrequent. The first must await the accumulation of a large number of bursts with well measured positions, easy only for the rarer large bursts which are recorded by several satellites. The second is in contrast with the situation for many random astronomical phenomena, for which the count rate is in some way inversely dependent on size, that is,

smaller bursts would be expected to be more frequent. Smaller bursts are indeed observed to be less frequent, but unfortunately just at a size where current satellites have difficulty in detecting them.

To resolve this, a number of balloon-borne experiments have been carried out with very large detectors. Such detectors, some hundred times larger than those at present carried in satellites, should be capable of seeing bursts about two orders of magnitude smaller in size. Their lack of confirmed observations has enabled some upper limits to be placed on small burst rates, with consequences for burst origins (for example Carter *et al.* *Nature* **262**, 370; 1976). However Cline *et al.*, at the Goddard Space Flight Centre, have revealed a source of confusion for such experiments (*Nature* **266**, 694; 1977). In an attempt to detect very small bursts, and to prove that any bursts seen were of cosmic origin, two similar detectors were flown on balloons separated by at least 1,400 km. They achieved 20 h of simultaneous observations, but detected no coincident bursts of cosmic gamma rays. Slow 'bursts' were observed on each individual payload and some were associated, but separated in time by many seconds. Whatever the cause of these (suggested to be due to magnetospheric activity), they constitute a serious source of background noise for burst experiments using a single balloon or near-Earth orbit satellite. As these authors point out, much more sensitive experiments are needed to enable small bursts to be seen and their expected anisotropy to be observed.

It may be that the theoretical interpretation of bursts may be helped more immediately by the observation of some correlated phenomenon. One particular burst was observed on 16 August last year by one of the chain of Vela satellites responsible for the original discovery, by two new satellites, Solrad 11A and 11B (Laros *et al.*, this issue of *Nature*, page 131) and by a transatlantic balloon experiment. The gamma-ray results published do not by themselves mark out the burst as very special (except as a simul-

taneous observation by a number of different experiments, including a balloon flight). However, what may be a very significant radio observation was made by Mandolesi *et al.* (*Nature* **266**, 427; 1977) at Medicina and Trieste. They observed on 16 August 1976 a radio burst at three widely spaced frequencies at Medicina, and at a fourth frequency using a 10-m parabola at Trieste. All lasted about 30 s, the duration of the gamma-ray burst, and they rule out man-made, atmospheric and solar effects as its cause. It occurred less than one minute after the gamma-ray burst, to which coincidence they ascribe a chance probability less than 10^{-4} . If the satellite burst position is eventually found to agree with the radio position, it will confirm the first such burst-correlated phenomenon. The separation in time would then be a consequence of the evolution of the burst source and would limit the models proposed to explain both bursts. Since this coincident radio burst was discovered only about a month after full operation of the radio array, and if the two bursts had a genuine common source, then one may hope that further coincidences will soon follow. □



A hundred years ago

THE WOODPECKER.—In the April session of the German Ornithological Society Prof. Alton concluded the recital of his investigations on the habits of the woodpecker. The peculiar drumming sound often caused by it was shown on various grounds to be entirely disconnected with the search for insects as hitherto supposed, and was regarded as a call to the opposite sex. Dr. Brehm defended the woodpeckers against the charge of seriously injuring the trees, and considered the slight damage resulting from them as more than compensated by the colour and animation which they gave to the otherwise sober and quiet forests.

From *Nature* **16**, 10 May, 30; 1877.

Sequencing DNA

from Maria Szekely

It is somewhat unusual that a new technique should already be being used all over the world at the time of its publication, but this is exactly what has happened with the ingenious new method of DNA sequencing worked out by A. Maxam and W. Gilbert at Harvard University. The method is now described in full detail (*Proc. natn. Acad. Sci. U.S.A.* **74**, 560; 1977), but already DNA sequences established in this way have been published—by Walz *et al.* (λ promoter-operator region (*Nature* **262**, 665; 1976)), by Fiers' group (SV40 DNA (*Cell* **9**, 117; 1976)), by Musso *et al.* (control region of the *gal* operon (*Proc. natn. Acad. Sci. U.S.A.* **74**, 106; 1977)), by Sures *et al.* (histone genes (*Cell* **9**, 495; 1976)) and others, showing how a useful technique can spread at amazing speed among scientists working in this field.

DNA sequencing, which some years ago seemed a formidable task and was approached by trying to convert it to an RNA sequencing problem (Gilbert and Maxam themselves sequenced the *lac* operator in 1973 (*Proc. natn. Acad. Sci. U.S.A.* **70**, 3581) by transcribing it into RNA and determining the sequence of this transcript) is today a relatively simple operation which can be accomplished by routine methods much more quickly than RNA sequencing could ever be done. In fact, the present trend is to convert RNA sequencing problems into DNA sequencing tasks to speed up the work.

Two very efficient techniques are available today to establish DNA sequences: Sanger's 'plus and minus' technique (*J. molec. Biol.* **94**, 441; 1975), based on enzymatic copying of single-stranded DNA, the spectacular results of which have recently been published in *Nature* (see **265**, 687; 1977) and Gilbert and Maxam's technique based on chemical modification and breakdown which has been applied to a variety of double-stranded DNA fragments. Both methods have in common a very convenient and fast way of interpreting the data: the sequence can be read off the gel electrophoretic pattern obtained from a complete mixture of different sized fragments. This readout technique makes it possible to establish long sequences in one experiment; the longest nucleotide stretches which can be analysed are in the range of 100–150 nucleotides and

the upper limit depends only on the resolution attainable in the gel electrophoresis.

The basic principles of the two methods are quite different however. Sanger, in the 'minus' technique, copies a single-stranded DNA stretch, synthesising different length fragments, each of which ends before a given base. Running four parallel samples in which stoppages occur at different bases provides sufficient information to read the sequences of the whole DNA fragment. The 'plus' technique serves to confirm the sequence; here the fragments are degraded by making use of the exonuclease activity of T4 DNA polymerase, under conditions when this degradation stops at sites where a given base is present. Again, four samples are run in parallel to determine the positions occupied by all four bases in the DNA fragment.

In the technique of Gilbert and Maxam no copying is involved, nor any enzymic reactions. Chemical reagents are used to modify and excise a given nucleotide from the polynucleotide chain. If conditions are mild, only partial modification occurs and only about one in 50 or 100 nucleotides is excised. In a random mixture there will be fragments of different lengths ending at all possible positions where this base was present. Four parallel samples are prepared also in this method to locate the positions of all four bases along the DNA chain.

In many cases either method can be used successfully for DNA sequence determination. The choice of the appropriate technique depends mainly on the single or double-stranded nature of the DNA. Sanger's technique requires single-stranded DNA whereas the Gilbert–Maxam procedure works equally well with single and double-stranded molecules. One variant of the Gilbert–Maxam procedure involves denaturation of the DNA fragment and separation of the strands, this step can, however, be omitted and another strategy used instead; the 5'-termini of a double-stranded DNA fragment can be labelled and the fragment split in the middle. This treatment yields two smaller fragments in each of which only one strand carries a labelled group. Consequently, only the products from this strand will show up on the autoradiographs. This represents a great advantage, as strand separation may prove difficult when applied to longer DNA molecules (although the Gilbert–Maxam separation method seems to work very well for restriction

fragments). As to the length of the DNA molecule to be sequenced, the Gilbert–Maxam technique has been mainly used so far for determining the sequence of restriction fragments and is probably best suited for fragments 100 to 200 nucleotides long. In the plus and minus technique, copies of any desired length can be synthesised from a convenient and well-defined start point, irrespective of the length of the DNA template. The method has also been used on the intact complete genome of ϕ X 174 phage (Sanger & Coulson *J. molec. Biol.* **94**, 441; 1975). The start point is determined by the oligonucleotide used for priming the synthesis of the labelled copy and can thus be chosen more or less at will.

Apart from the very great progress in elucidating the structure of different genomes, these techniques may well prove of great advantage in the field of RNA sequencing. Complementary DNAs have already been produced for sequence studies on eukaryotic mRNAs (see for example Proudfoot *J. molec. Biol.* **107**, 491; 733; 1976 and Cheng *et al. ibid.* 527). These cDNAs may now be accessible to the new sequencing techniques. We may expect in the near future a great advance in the sequence analysis of this group of mRNAs which are of great functional importance but which so far have presented serious problems for sequence studies. □

Man-made auroral pulsations

from Alan Johnstone

PULSATING auroras are one of the most intriguing types of auroral display. Irregularly shaped patches of light appear to turn on and off regularly with a period of the order of 10 s. The beam of electrons, which produces the auroral light when it strikes the atmosphere, must be modulated by some plasma wave, on its way down the magnetic field line from the outer radiation belt. Despite much speculation about the cause no satisfactory explanation has been found. The report

Maria Szekely is a lecturer in the Department of Biochemistry at Imperial College London.

Alan Johnstone is at the Mullard Space Science Laboratory at University College London.

AN important control on the immune response in many animals, including rodents and primates, appears to be provided by the immune response (Ir) genes which are tightly linked to the major histocompatibility complex (MHC). These genes are antigen specific in their effects, that is, they appear to play an important part in determining whether an animal can respond to any particular antigen, but as yet their function and site of action is still unclear. However, evidence is rapidly accumulating that the Ir genes are in fact expressed in macrophages or the macrophage-like cells that process or present antigen to T lymphocytes.

Rosenthal and Shevach (*J. exp. Med.* **138**, 1194, 1213; 1973) showed that guinea pig T cells from an F₁ hybrid between a responder and a non-responder could only be stimulated to proliferate by antigen-coated responder (or F₁) macrophages, but not by non-responder macrophages. Subsequently, analogous proliferative studies have been performed in the mouse by Schwartz and his colleagues (*J. Immun.* **117**, 531; 1976 and *Proceedings 3rd Ir Gene Workshop*, Asilomar, 1976). These studies suggested that the proliferating cells recognised antigen in association with macrophage Ir gene products, which since they were inhibited by anti-Ia antisera (antisera directed against the I region-associated antigens of the mouse MHC) were probably macrophage Ia antigens. While the immunological role of the proliferating cell(s) in these assays is not known, analogous results have been obtained in assays for the induction of antibody (Pierce *et al. J. exp. Med.* **144**, 371; 1976) or of helper cells (the T cells which collaborate with B cells to enable the latter to produce antibody). For example, Erb *et al. (Eur. J. Immun.* **6**, 365, 1976) have demonstrated the capacity of macrophage Ia

Marc Feldmann is an ICRF research fellow working at University College London.

Are the Ir genes expressed by macrophages?

from Marc Feldmann

antigen, complexed to a fragment of immunogen, to form a soluble complex of molecular weight ~55,000 daltons, which replaces intact macrophages in the induction of helper cells. All these studies reinforce the concept that there are Ir gene effects expressed in macrophages. J. E. A. P. Miller and his colleagues have recently obtained compatible results in a delayed hypersensitivity (DTH) model in the mouse, leading to the same conclusion for the Ir genes involved in DTH (Miller *et al. J. exp. Med.*, in the press).

In this issue of *Nature*, Rosenthal, Barcinski and Blake (page 156) extend this point. Using insulin molecules from various species, with minor differences in amino acid sequence, they demonstrate Ir gene control of antigen recognition at the level of the macrophage, measured by the proliferative response of guinea pig T cells. Their results are most easily interpreted by postulating that macrophages may select the appropriate portion of the antigen to be presented to T cells. How is this done? The simplest possibility is that macrophage Ir genes define a class of receptors of broad specificity that 'focus' or orient antigen for T cells to recognise it. Conceivably, these receptors may be the descendants of the recognition molecules which antedate the vertebrate immune system. If that interpretation is correct the antigen-presenting cells, and not the T cells, dictate the antigen specificity of the Ir gene effect. But since T cell responses depend so intimately on macrophage presentation, macrophage Ir genes would reveal them-

selves (as has been found) as defects of T cell function. The search for the immune response gene products may then already be partly over—they have been isolated and partly purified as the macrophage Ia antigen-immunogen complex, which induces the activation of helper cells *in vitro*.

The idea that the Ir genes are expressed on macrophages is also compatible with other findings that have been interpreted quite differently. Munro and Taussig for example (*Nature* **256**, 103; 1975), described two types of mice with genetically determined low responses to the synthetic polypeptide (T,G)-A-L; those which did not make T helper factor (an antigen-specific mediator of T cell help) and which was interpreted as a T cell defect, and those which did not respond to T cell helper factor, which was interpreted as a B cell defect. But since macrophages are needed for both these reactions a macrophage defect could be at the root of both these failures. T cell suppression has also been proposed as a mechanism of non-responsiveness and since suppression is much less dependent on macrophage function than T cell help, a net suppression would occur in the presence of dysfunctional macrophages. This idea is not incompatible with the proposed role of the macrophage.

Whatever the truth of the mechanism of Ir gene action, it now seems certain that macrophages and macrophage-like cells have a more complex role in the immune system than hitherto suspected, and the precise nature of the antigen-presenting cell needs to be re-examined. Are these classical macrophage or lymphocyte-like cells, or a hybrid between the two. In view of the plethora of lymphocyte subpopulations recognised in the past two years, the existence of yet another type of cell in the immune system need not be too improbable. However, the question still remains as to whether there are Ir genes expressed by other cells of the immune system.

that pulsations have been induced artificially near a pulsating display is unexpected and might give valuable insight into the mechanism causing the natural phenomenon.

The artificial pulsations were found by Deehr and Romick (this issue of *Nature*, page 135) in one of the natural auroral spectral lines at 200 km altitude, just below a barium release in the F region of the ionosphere. This technique, in which 1.5 kg of Ba is ejected by a thermite reaction from a canister carried aloft by a rocket, is used to measure electric fields and

winds in the upper atmosphere. The barium cloud is partially ionised by sunlight, forming a cloud of ions which resonantly scatters purple light from the Sun and a neutral cloud which scatters green light. The neutral cloud is blown across the sky by winds while the ionised cloud is driven by the combination of electric fields and winds. By tracking both clouds optically from the ground the effects of the two forces can be separated. The technique is very effective and has been used many times in the last decade.

Ideally the barium should not dis-

turb the medium it purports to measure. Small barium releases, such as that used by Deehr and Romick, do not distort the electric fields seriously although they must have some effect. Their observation is not the first of perturbations in the ionosphere beneath a barium cloud. Stoffregen (*J. atmos. terrest. Phys.* **32**, 171; 1970) detected a temporary increase in the auroral line at 557.7 nm at 100 km altitude below a barium cloud. He did not observe pulsations and there was apparently little natural aurora present. Spracklen and Jones (*Q. Jl R. Astr.*

Soc. **17**, 469; 1976) have observed an increase in E region (100 km) electron densities by radio techniques. Both authors have attributed this coupling between the F region where the barium was released, and the E region where the disturbance is observed to the flow of charge from one side of the cloud to the other to neutralise the build up of a polarisation electric field as the cloud starts to move. The horizontal current, across the magnetic field, can only flow in the E region where there is sufficient Pedersen conductivity to allow it. The downward flow of electrons from one side of the cloud enhances the electron density in the E region while the upward flow on the other side should leave a lower density.

The release reported by Deehr and Romick is the only one to take place within a significant amount of natural aurora and its presence seems to be a necessary condition for the pulsations. They were observed at 200 km where electrons of 500 eV deposit most of their energy. Higher energy electrons would penetrate further into the E region, but any effects there might have been were masked by natural aurora.

Natural pulsating auroras are found

predominantly at altitudes below 100 km (Brown *et al. Geophys. Res. Lett.* **3**, 403; 1976) implying that higher energy electrons, between 10 keV and 60 keV, are usually involved. Bryant *et al. (Planet. Space Sci.* **23**, 867; 1975) measured the electron fluxes producing pulsating aurora directly and also found relatively high energy particles. Further, they were able to detect phase differences between the pulsations observed at different energies, with the higher energy pulsations seen first. This velocity dispersion placed the source of the modulation more than 40,000 km away, and certainly not within 300 km.

The differences between what is known of natural pulsations and the circumstances surrounding the observations of the artificial ones seem to suggest that the existence of the two displays side by side was coincidental. It would be hasty to draw such a conclusion when so few hard facts are known about either phenomenon and when the range of possible mechanisms is still so wide. One thing is certain—there will be determined attempts to repeat the observation during future barium release experiments wherever they are performed. □

It seems likely that complex interactions between peptidergic neurones will be uncovered since neurones originating in dorsal root ganglia and separately containing Substance P and somatostatin enter the spinal cord where they overlap with neurones containing enkephalin, neurotensin and TRH. Peptide-monoamine interactions also seem likely in view of the high density of enkephalin- and Substance P-containing neurones in areas containing cell bodies for serotonin, dopamine and noradrenaline. Iversen reported experiments that indicate an inhibitory control of Substance P release by GABA, and raised the possibility that these putative transmitters represent excitatory and inhibitory controls respectively for dopaminergic mechanisms in the basal ganglia. This may have important consequences in our understanding of various disease states such as Huntington's chorea and Parkinson's disease.

New results and controversy still abound in the opioid peptide field. The potent analgesic and behavioural effects of β -endorphin are the subject of considerable research (D. de Wied (University of Utrecht) and D. G. Smyth (National Institute for Medical Research, London)) and now it seems that stable analogues of enkephalin produce similar effects, even on intravenous administration, according to reports by R. Miller (Wellcome Foundation) and A. Z. Rónai (Budapest). Unfortunately there is little evidence yet that these short chain peptide analogues will prove to be less addictive than morphine. A promising note was sounded by L. Terenius (University of Uppsala) who described the possible anti-schizophrenic action of the opioid antagonist naloxone in Swedish patients, thus implicating endorphins in some mental disease state. Caution was sounded here, however, by the report that American psychiatrists had failed to reproduce these results. Terenius also finds increased levels of endorphins in the cerebrospinal fluid of schizophrenic patients and results in this area may prove easier to replicate than in the more problematic area of the clinical assessment of schizophrenic remissions. A further link between analgesia and endorphins was reported by Terenius and J. Hughes (University of Aberdeen) who both found increased opioid peptide levels in the cerebrospinal fluid of patients receiving intracerebral brain stimulation for the relief of chronic pain.

The major impression left by this symposium was the impressive accumulation of evidence for central peptidergic transmission but the lack of evidence relating to defined physiological roles. However, most par-

Centrally active peptides

from John Hughes

A symposium on Centrally Active Peptides, organised by the Biological Council, was held at the Middlesex Hospital Medical School, London, on 4-5 April, 1977.

AN extra dimension has been added to neurobiological research in the past few years with the recognition of the potential importance of peptide-releasing neurones in the brain. The concept of peptidergic neurotransmission and neuromodulation has reached fruition as an area of multidisciplinary research which promises major advances in our understanding of normal and abnormal brain function.

Important advances in the distribution of peptides were reported and the great value of immunohistochemical techniques in this field was emphasised. A picture emerged of an extensive distribution of peptidergic neurones throughout the brain and spinal cord. Prominent amongst these are neurones containing Substance P and enkephalins; these systems seem to rival and perhaps exceed that of the better known catecholamines. A significant advance has been the demonstration of

cell bodies for peptidergic neurones (R. P. Elde (University of Minnesota), S. H. Snyder (Johns Hopkins University) and L. L. Iversen (MRC Neurochemical Pharmacology Unit, Cambridge)). It seems that the enkephalins are contained in short interneurones with a terminal distribution closely paralleling that of opiate receptor binding sites. This system appears to impinge on areas receiving primary sensory inputs as well as on those involved in the control of respiratory, cardio-vascular, somatic motor and neuroendocrine functions.

Iversen presented cogent evidence for a central neurotransmitter role for Substance P, and this peptide is now a firm favourite as one transmitter involved in the relay of information by small primary afferent sensory fibres. Multiple transmitter roles for Substance P also seem likely from its distribution in other areas such as the basal ganglia, hypothalamus and limbic system. Other neuronal systems containing vasopressin, oxytocin, vasoactive intestinal peptide, thyrotropin releasing hormone (TRH), somatostatin and neurotensin have also been identified in many brain areas (Elde, M. J. Brownstein (National Institute of Mental Health, Bethesda), A. G. E. Pearse (Royal Postgraduate Medical School, London) and Snyder).

John Hughes is Deputy Director of the Unit for Research on Addictive Drugs at the University of Aberdeen.

ticipants seemed confident of finding further evidence for the participation of peptides in normal and abnormal brain function. Continuing research is likely to lead to a deeper understanding of neuronal physiology with corresponding therapeutic benefits. □

Membranes and receptors

by Peter Newmark

A conference on Receptor-Signal Transmission was held in Titisee on 30 March–3 April, 1977. It was sponsored by the Dr Karl Thomae GmbH, a subsidiary of Boehringer-Ingelheim.

A CLUSTER of receptorologists, convened for three days in the Black Forest, pondered not only the mobility of receptors but also the transmembrane activation of adenylate cyclase, mechanisms of cell adhesion and membrane changes in secretion.

The presentation of a ligand to its receptor may be an event of considerable complexity involving cooperative effects whereby the initial binding can either promote or hinder further binding. One of the best examples of this is the negative cooperativity demonstrated in the binding of insulin to its cell surface receptors. P. de Meyts (National Institutes of Health, Bethesda) was cautious in his discussion of the mechanisms of negative cooperativity. He has however obtained evidence that the avian insulin receptor still exhibits negative cooperativity when solubilised. The solubilised receptor exists in high molecular weight form in the absence of insulin, and is converted, reversibly, to low molecular weight form by addition of insulin. From that de Meyts suggests that negative cooperativity, often supposed to be the result of clustering of receptors, instead involves the dissociation of a high affinity tetramer form of the receptor to a low affinity monomer. de Meyts has also been looking for negative cooperativity with a series of insulin analogues and the results, interpreted in conjunction with the known three-dimensional structure of insulin, indicate that cooperativity is determined by a discrete domain on the surface of the insulin molecule distinct from that involved in binding.

A highlight of the meeting was the general, probably unprecedented, agreement that emerged from the various contributions on the molecular inter-

actions involved in the transmembrane activation of adenylate cyclase following the interaction of ligand with receptor (M. Rodbell, NIH, A. Levitzky, University of Jerusalem, E. Helmreich, University of Wurzburg and M. Schramm, University of Jerusalem). All agreed that the receptor is not physically coupled to the adenylate cyclase moiety and that the activation of the enzyme as a consequence of the ligand interacting with its receptor is a GTP-dependent process. The GTP, whose origin is intracellular, is thought by Rodbell to activate both the receptor and the adenylate kinase each by way of a separate nucleotide binding site. Levitzky favours the view that only the enzyme is activated by GTP. That the GTP activation is normally transient and is terminated by hydrolysis of the nucleotide is shown by the permanent switching on of adenylate cyclase by GTP analogues that resist hydrolytic degradation. Some progress was reported in the purification of the relevant GTP binding proteins (Rodbell and Helmreich) and of both the β -adrenergic (Levitzky) and epidermal growth factor receptors (C. F. Fox, University of California, San Francisco).

From the general agreement that

there was no direct, permanent association between receptor and enzyme, it can be predicted that a receptor from one cell should be able to couple functionally with the adenylate cyclase from another cell when the two cells are fused. Schramm reported experiments which bore out that prediction. In initial experiments (*Proc. natn. Acad. Sci. U.S.A.* **73**, 4410; 1976) Friend leukaemia cells, which do not possess β -adrenergic receptors, were fused with turkey erythrocytes which had been pretreated with N-ethylmaleimide or heat to inactivate their adenylate cyclase (not their receptor as mistakenly stated in *Nature* **265**, 681; 1977). Ghosts of the fused cells exhibited a low basal adenylate cyclase activity that was markedly activated by the β -adrenergic drug isoproterenol. Since then equivalent results have been obtained with mixtures of quite different cell types. For example the β -adrenergic receptor of rat C6-glioma cells can be coupled with the adenylate cyclase of mouse Y-1 adrenal tumour cells. What is more the affinity of the receptors for various agonists is unaltered after being coupled to a foreign adenylate cyclase.

I. Pastan (NIH) reported a series of studies that have led him to the view

Multinucleon transfer reactions

from P. E. Hodgson

MULTINUCLEON transfer reactions are now used extensively to study nuclear structure, and recently a group at Los Alamos has found remarkable similarities between the spectra of states excited in the same final nucleus by the ($^6\text{Li},d$) reaction on the one hand and by the (t,p) and (h,n) reactions on the other (*Phys. Rev. Lett.* **38**, 587; 1977).

The group measured the ($^6\text{Li},d$) reaction on ^{54}Fe , ^{56}Fe , and ^{58}Fe for excitation energies up to about 4 MeV in the final nuclei, using a lithium beam of 34 MeV. About 20 states are known in this region of excitation for these nuclei, and it was found that only a few of them are excited by this reaction, and with few exceptions these are just the states also excited by (t,p) reactions going to the same final nuclei.

The exceptions are the O^+ states at about 3.5 MeV in each of the final nuclei that are excited by ($^6\text{Li},d$) but not by (t,p). These are very probably the states strongly excited by the (h,n) reaction, and interpreted as two-particle two-hole pairing vibration states in the $Z=28$ proton closed shell.

The group therefore concludes that the ($^6\text{Li},d$) reaction is sensitive to the same neutron and proton pairing correlations that govern the correspond-

ing two-nucleon transfer reactions. This conclusion supports recent work on the alpha-particle spectroscopic factors, in which the four-particle amplitude is factored into coupled two-nucleon structure factors. The similarity between the ($^6\text{Li},d$) and (t,p) reactions is then explained as the transfer of the proton pair to the ground state and the neutron pair to the ground and excited states as in the (t,p) reactions. The similarity between the ($^6\text{Li},d$) and (h,n) reactions is explained in the same way.

This work suggests further experiments on the two- and four-particle transfer reactions to determine the spectroscopic factors and to relate them to the details of nuclear structure. It also shows that the ($^6\text{Li},d$) reaction can be used to study two-particle excitations that cannot be studied directly by two-particle transfer reactions because the requisite target nucleus is unstable. Such studies should also shed further light on the alpha-particle spectroscopic factors that are important for the study of alpha-clustering in nuclei.

P. E. Hodgson is in the Department of Nuclear Physics at the University of Oxford.

Peter Newmark is the Biological Sciences Editor of Nature.

has evidence of an asymmetric distribution of a possible molecular recognition system.

Is this the first toe-hold in putting the process of retino-tectal pattern formation on a firm biochemical footing? Or is it a futile exercise that cannot hope to explain the complexities involved? We all went home! □

Positive ions in the stratosphere

from L. Thomas

OBSERVATIONS of ion composition with rocket and satellite-borne mass spectrometers have made a major contribution to our understanding of the ionised regions of the upper atmosphere but have been restricted to heights above 60 km because of experimental difficulties. In the lower part of the height range the relatively high ambient pressures make it necessary to use a high speed pump for maintaining the spectrometer within its operating pressure range, and the small ion concentrations give low signal currents. Another complication encountered in the rocket observations at these lower heights has been the fragmentation of weakly bound ions through thermodynamic break up at the shock front associated with the supersonic velocity and in collisions, because of the electric fields required to draw ions into the sampling orifice. The difficulties are accentuated at still lower heights because of the greater atmospheric pressures and the increased rocket velocity. Consequently, for stratospheric levels, information is available only from ion mobility measurements but these cannot be used to identify the different ion species present.

F. Arnold, D. Krankowsky and K. H. Marien of the Max-Planck-Institut für Kernphysik, Heidelberg, have developed a quadrupole mass spectrometer with a cryogenically cooled pump, an improved sampling procedure, and a relatively low draw-in potential to reduce the effects of ion fragmentation and, thereby, have managed to extend ion composition measurements down into the stratosphere. They have reported three sets of rocket measurements over the height range 35–55 km (*Nature* **267**, 30; 1977). Although the mass resolution was limited to 2 a.m.u. to ensure sufficient sensitivity, the authors have been able to make a tentative identification of the ions registered. Above

about 40 km the most abundant ions were found to be consistent with proton hydrates, $H^+(H_2O)_n$, with n ranging from 1 to 3. It has been known for more than a decade that such hydrates dominate the positive ion composition over the height range 60–85 km in normal conditions, although some doubt still exists about the appropriate value of n in the ambient ions. Below 40 km, the present results show that a marked change in the predominant ion species occurred within a height range of a few kilometres, and the authors have interpreted the ion masses as resulting from reactions of the proton hydrates with minor neutral species; these species are required to have a higher permanent dipole moment or proton affinity than water to permit the replacement of water molecules or proton transfer to occur. On the basis of aeronautical arguments and supporting laboratory data Arnold *et al.* propose formaldehyde as the neutral species involved. This is known to be a major product of the oxidation of methane, a constituent whose height distribution is reasonably well established. The concentrations of formaldehyde estimated from the interpretation of the ion composition measurements are consistent with those derived in photochemical models (see for example Stewart & Hoffert *J. atmos. Sci.* **32**, 195; 1975). It seems likely that fragmentation has occurred in the ion sampling and, consequently, the particular proton hydrates and formaldehyde-based ions are probably not truly representative of stratospheric conditions. However, as the authors point out, it might be expected that the fragmentation will have removed water and formaldehyde molecules without modifying the central ions.

At stratospheric heights ionisation is produced by the interaction of galactic cosmic rays with atmospheric gases, the primary ions produced being those of molecular oxygen and nitrogen. The formation of proton hydrates from these ions is well understood from laboratory measurements and numerical studies of the ionospheric D region (Ferguson *Rev. Geophys. Space Phys.* **9**, 997; 1971). It would be expected that the ion chemistry of the stratosphere would involve minor neutral species in addition to water, because of the greater variety and concentrations of these species at these heights. In this connection, the inference concerning the upper limit on ammonia concentrations drawn from the absence of relevant ions is of interest. The value quoted is rather smaller than would be expected from loss by photolysis alone (McConnell *J. geophys. Res.* **78**, 7812; 1973) and could indicate the rapid loss of ammonia by collisions with, or con-

version to, aerosol particles.

Considerable interest is currently being shown in the chemistry of the stratosphere. This has arisen out of concern about the catalytic depletion of ozone by chlorine atoms produced by the photodissociation of chlorofluoromethanes (CFMs) transported from the Earth's surface (Rowland & Molina *Rev. Geophys. Space Phys.* **13**, 1; 1975). The observations of ion composition are relevant to the question of whether ion-molecule reactions might play a significant part in the loss of the chlorine-bearing compounds. At the height of the observations reported, the CFMs will be photodissociated and attention needs to be concentrated on the photodissociation products, namely chlorine atoms, chlorine oxide and hydrogen chloride. No exothermic reactions are known to occur between proton hydrates and these three gases (Fehsenfeld *et al. J. geophys. Res.* **81**, 4454; 1976) but the corresponding information relating to the ions identified below 40 km is not available.

Apart from the immediate considerations of environmental protection, the study of the stratosphere has become an active area of atmospheric research. It seems certain that mass spectrometer techniques for the measurement of ion composition will be exploited both in rocket and high-altitude balloon flights. The measurements will serve to complement the observations of neutral minor constituents which are being pursued vigorously and are also likely to inspire further studies of ion-molecule reactions. □

Whither protein structures?

from S. D. Dover

THE pipe-dream of designing and producing a protein to carry out a specific task has been brought nearer reality by the advent of recombinant DNA which offers the possibility of isolating, mutating and cloning particular genes. The resulting change in amino acid sequence specified could produce the intended structural change in the protein. But this will only be possible with an adequate theory of the effect of sequence changes on structure, based firmly on experimental structural information. Without this theory at the level of atomic detail,

S. D. Dover is a lecturer in the Department of Biophysics at Kings College London.

L. Thomas is a member of the staff at the SRC Appleton Laboratory.

designing new structures would be a hit-or-miss affair.

Interest in the solution of the problem of analysis and, subsequently, prediction of protein structure from sequence has an extensive history. From the studies of Ramachandran (*J. molec. Biol.* **7**, 95; 1963), to those of Levitt (*Nature* **261**, 552; 1976; *J. molec. Biol.* **104**, 59; 1976) theoretically oriented work has been progressing. The basic data for these studies have been the protein structures which regularly appear in the literature. The problem has tended to be approached from the point of view of overall structure with atomic details being left until later.

This has been partly due to the fact that most protein structures available are unrelated. Of the related structures the globins are probably the ones likely to reveal the greatest comparative information. The overall picture of haemoglobin and its allostery has been described by Perutz. Not only is there an abundance of species and within-species mutant haemoglobins of the four-chain type, but also single chain oxygen carriers such as haemerythrin (*Proc. natn. Acad. Sci. U.S.A.* **72**, 2160; 1975), erythrocrutorin (*Eur. J. Biochem.* **19**, 42; 1971) and the monomer/dimer system of the sea lamprey (*J. molec. Biol.* **74**, 331; 1973). Consideration of the chain interactions in atomic detail is likely to be stimulated by recent refinement of sperm whale myoglobin.

Takano (*J. molec. Biol.* **110**, 533, 569; 1977) has just published the first refinement of the deoxy- and met-structures since Kendrew *et al.* (*Nature* **190**, 666; 1961) first interpreted their map. Using new diffractometer data and variants of the well-tried method of alternating phase and realspace refinement (*Acta Cryst.* **A27**, 436; 1971) structures suitable for detailed comparison with haemoglobin were achieved. Takano's stereo diagrams show rather clearly the similarity of structure despite the differences in sequence. The possibility of extensive comparisons of this type is most encouraging, but it does raise a number of problems. Should one compare only whole myoglobin molecules or also separated sections of the chain? How can one objectively make the comparisons between non-identical objects and quantify and display the comparisons?

The relationship between haemoglobin and myoglobin provides information at all levels, but the relative changes are greater than one might want for truly secondary structure. As a starting point movements smaller than those produced by a single amino acid change would be interesting. Such movements would be expected to occur immediately before denaturation, even

though it is a cooperative phenomenon. Stability studies on haemoglobin have already been started (see *Nature* **247**, 341; 1974; *Nature* **255**, 256; 1975). Extension to full crystallographic refinements with crystals at, for example, different pH and ionic strengths might be the key to a number of problems associated with small atomic movements. But will there be PhD students to do all those structures? □

Heat flow at the mantle surface

from Peter J. Smith

SOME time ago Chapman and Pollack (*Earth planet. Sci. Lett.* **28**, 23; 1975) took what they called "a new look" at the Earth's thermal characteristics, producing in the process the first worthwhile map of the global surface heat flow. Then, as now, there were still large areas of the world with few or no data; but Chapman and Pollack had hit upon the idea of filling these gaps with heat flow values predicted from known empirical relationships between heat flow and tectonic age. Both real and predicted observations were then incorporated into a 12th degree spherical harmonic analysis which was translated into a global contour map.

Pollack and Chapman (*Earth planet. Sci. Lett.* **34**, 174; 1977) have now extended their analysis to the much more difficult problem of plotting the heat flow at the crust-mantle boundary. Mantle heat flow is, of course, the total heat flow less that produced or released in the crust; but it is not easy to determine just what the crustal contribution is. Most heat produced in the continental crust comes from the decay of radioactive isotopes whose concentration decreases with depth in some only vaguely understood way. Most heat from the oceanic crust, on the other hand, is released as the lithosphere produced at oceanic ridges cools, although there is also a small but significant radiogenic component.

Combining theory with experimental results, Pollack and Chapman nevertheless manage to estimate these crustal contributions to the total heat flow, whence they derive sufficient data to carry out an 18th degree spherical harmonic analysis. The resulting global contour map of heat flow at the mantle

surface differs from the corresponding crustal surface map in two notable respects. First, a comparison makes it clear that the near equality of continental and oceanic heat flows at the Earth's surface does not extend down to the mantle; at the crust-mantle boundary the mean heat flow in continental regions is only about half of that in oceanic regions (28 mW m^{-2} and 57 mW m^{-2} , respectively). This means that the mantle below the oceans must be considerably hotter than that beneath the continents as several theoretical models have already suggested (see for example, MacDonald *Rev. Geophys. Space Phys.* **1**, 587; 1963).

The second result is that there is much less heat flow contrast within continental regions at the mantle surface than there is at the crustal surface. In Australia, for example, the mean crustal surface heat flow is 81 mW m^{-2} in the Proterozoic rocks of the central region and 39 mW m^{-2} in the western shield. At the mantle level both regions have heat flows of 27 mW m^{-2} . Evidently within-continent heat flow variations at the Earth's surface are due entirely, or almost entirely, to regional differences in crustal radioactivity and/or erosion. At the crust-mantle boundary the heat flow contours are simple closed curves centred on the continents and showing no regional complexities.

Finally, Pollack and Chapman are able to demonstrate correlations between the heat flow fields and other geophysical parameters. For example, both total and mantle heat flows correlate positively with the Earth's gravitational potential. However, the mantle heat flow correlates the less positively, suggesting that the broad geographic variations in the gravity field may be due in part to thermally produced heterogeneities in the crust. On the other hand, the fact that the positive correlation remains when the crustal contribution is removed suggests that thermal heterogeneities extend well into the mantle.

By contrast, both the total and mantle heat flows correlate negatively (the latter the more so) with topography, although this is produced by the first-order topography of the Earth (the gross difference in levels between continental and oceanic crust) which is largely a petrological rather than a thermal effect. When the continental and oceanic levels are normalised to remove this petrologically produced contrast, the total and mantle heat flows are found to correlate positively with topography, emphasising, according to Pollack and Chapman, "the significant role that temperature plays in the morphology of both continents and oceans." □

Peter J. Smith is in the Department of Earth Sciences at the Open University.

review article

Economics of alternative energy sources

Martin Ryle*

An important part of the oil and natural gas at present consumed in the UK is used for the heating of buildings, a demand which shows large diurnal, day-to-day and annual fluctuations. The replacement of this energy by nuclear-generated electricity, as at present envisaged, would require the construction of some 250 GW of additional capacity by the end of the century, a programme which does not seem feasible. By incorporating relatively cheap, short term storage in the form of low-grade heat, the generating capacity required to fulfil peak demand could be reduced by more than 50%. As soon as such storage is provided, however, other sources of energy become viable and attractive alternatives, and the UK is well situated to make use of wind, wave, and tidal power. It seems likely that the value of North Sea oil/gas reserves as feedstock to the chemical industry will rise sufficiently to make an early reduction in their consumption as fuel of great economic importance.

It is generally accepted that the energy at present provided by oil and natural gas will be approaching exhaustion by the end of the century, and that the resulting deficit must be made good by the widespread adoption of nuclear power, a large expansion of the coal industry or the development of alternative 'renewable' sources. Each of these possibilities has disadvantages and limitations, and there are problems of development on a short enough time-scale. Because the value of oil and gas as feedstock to the petro-chemical industries (which at present consume about 20% of the total¹) will increase rapidly as the reserves are depleted, it is of the greatest importance to introduce replacements as early as possible; each year gained represents a 4-year extension of supplies of raw material.

Since all of the potentially most useful renewable sources of energy in the UK (solar energy to supply heat or photo-voltaic electricity, wind, wave or tidal power) provide an intermittent or variable supply, their economic viability has always been based on whether they can be justified as 'fuel savers', and it has been assumed that they can never contribute to the 'base-load' in the sense of reducing the total installed capacity of conventional power stations. It then follows that, in order to be cost-effective, the interest on the capital costs of an alternative power system, together with its running costs, must be less than the fuel and running costs of a conventional system, without regard to the capital cost of the latter. Account must, of course, be taken of future inflation, interest rates, and relative costs of fuel in relation to the inflation rate, and the results of such analyses are often unreliable because of uncertainties about these factors, rather than in estimates of the engineering costs.

If the demand for power were constant, this treatment of alternative sources simply as fuel-savers would be a reasonable way of determining the viability of any particular system, but in practice this is far from being the case; there are very large variations in the demand for energy (diurnally, over the periods of 5-10 days typical of UK weather variations, and annually). At present, a com-

paratively small proportion (~ 15%) of the total useful energy is provided by electricity, and the variations in demand for it can normally be covered by the use of pumped hydro-electric stations, gas-turbine generators with a short run-up time, and by arrangements for selling off-peak electricity for water-heating and storage heaters. The fluctuating loads associated with space-heating of buildings and other consumers of oil and natural gas involve peak loads some 3 times the mean energy; at present short-term oil-storage can easily be supplied locally, with longer-term fluctuations covered by depôts or by varying the rate of abstraction through the year. The problem of accommodating the fluctuating loads when these latter energy sources are replaced by electricity does not seem to have been adequately discussed but would, in fact, be acute. I shall seek to demonstrate that a solution of this problem dramatically alters the balance in favour of 'renewable' sources of energy.

In the following sections an estimate is made of the distribution of oil and gas supplies amongst different users; the consequent diurnal, weather-system and annual variations of this additional electricity demand are then discussed in relation to the problems of generation and storage.

The figures are derived from UK Energy Statistics, 1973-75 (ref. 1) two NEDO reports on energy^{2,3}, the conference on 'Energy in the 1980s' organised by the Royal Society⁴, the Building Research Establishment Report on energy consumption in buildings⁵ and other sources. All energy values are expressed in terms of the electrical energy needed to replace them, so that they take account of the efficiency for the consumer of oil- or gas-burning heaters⁶; efficiencies of generation are taken into account in the establishment of the costs of different energy sources.

It is likely that for economic and other reasons the energy consumption in the UK will not increase in the manner assumed in earlier analyses²⁻⁴; better thermal insulation and the economic pressures towards lower temperatures in housing and other buildings, the production of longer-life goods, and better utilisation of industrial waste heat will all tend to maintain the recent trend towards lower energy

*Cavendish Laboratory, Madingley Road, Cambridge, UK.

consumption. For this reason it is not unreasonable to adopt the 1972–73 values for the energy consumption, the latest for which a detailed analysis is available³; the relative costs of different generating systems to replace the 1972–73 oil and gas consumption will in any case not be seriously affected even if the total varies from that of 1972–73. Greater uncertainties are caused by the rate at which coal production could (or should) be expanded, the development of more efficient methods of combustion⁶, and the relative costs of manufacturing gas, liquid fuels or other synthetic materials from coal instead of natural oil or gas, especially in relation to fuels for transport.

Although energy data for 1972–73 are used, the capital and operating costs are given in 1975 US \$ prices because much of the most useful data has been presented in this form; an annual inflation rate of 10% has been assumed for 1974–76.

Time distribution of additional requirements

It is assumed that the present (1972–73) UK electricity production (0.77×10^9 GJ p.a.) will continue to supply the present electricity demand, with the exception of the annual contribution (0.24×10^9 GJ) now provided by oil-powered gas turbines. Since the latter are used for peak load boosting, particularly in cold weather, it is assumed that their replacement will require an additional 0.024×10^9 GJ during the summer quarter, 0.06×10^9 GJ during each of the spring and autumn quarters and 0.096×10^9 GJ during the winter quarter; it is assumed that 2/3 of the latter will be needed during the cold spells of the 5–10-day weather pattern.

By taking the present distribution of oil and natural gas between each class of consumer⁵ we can derive the equivalent electrical energy demand and its variation with time of day, with outside temperature, and annually. The conversion efficiency of oil and gas heaters is taken as 60% (ref. 5) and that of electrical heating as 100%, although in some applications, such as water-heating for which off-peak use can help to even out the day-to-night demand, storage is necessary and an efficiency of 95% has been assumed.

The annual variation of the demand for space-heating in domestic and other buildings has been assumed, for simplicity, to be distributed uniformly throughout each quarter in the ratios 50%, 25%, 0% and 25%; this probably gives a rather low figure for the mid-winter requirements.

With present standards of building insulation, night/day switching is essential if a large increase in total demand is to be avoided, and a 50% cycle is assumed. Since the heat loss from a building is roughly proportional to the difference between internal and external temperatures (though it also depends on wind and other factors), there are large fluctuations with the timescale typical of UK weather patterns (5–10 days). A range of winter temperatures of -3 to $+11$ °C is assumed, with average 'whole-house' inside temperatures for the bulk of existing houses of 15.5 °C (60 °F). The hot-water demand is assumed not to vary significantly either annually or with weather conditions, and since 24-h storage is relatively simple it is assumed that two-thirds of the total demand could be restricted to off-peak times.

In the case of Industry it is assumed that, of the total equivalent electrical demand of 0.92×10^9 GJ p.a. at present supplied by oil and gas, there is a constant annual load of 0.72×10^9 GJ, of which 70% is consumed during the day, with an annually varying component of 0.2×10^9 GJ associated with space heating.

Some 95% of the transport energy is associated with road traffic, 67% of which is passenger traffic^{2,5}. Taking a weighted mean conversion efficiency between petrol (14%) and diesel (21%)² and a likely overall efficiency of 60%

for rechargeable batteries, hydrogen fuel cells or other possible storage methods, the equivalent electrical load is computed, with a constant annual rate, but on the assumption that 75% could be restricted to off-peak times in order to assist in reducing the day-to-night fluctuations.

The results are given in Table 1; the magnitude of the total energy demand, 2.1×10^9 GJ, should be compared with the present entire electrical output of 0.77×10^9 GJ, while it is evident that the fluctuations of demand require an extremely large additional installed capacity.

It has been suggested that a major increase in coal production might allow an increase of up to 50% by the end of the century. There are many alternative ways in which this increase might be distributed, but it is clear that part must be used to replace the present oil and natural-gas feedstock to the chemical industry. Assuming that this could be achieved with an efficiency of 90% of that of conversion from oil/gas, this would leave¹ an annual surplus of 0.36×10^9 GJ of equivalent electrical energy at 60% efficiency. This might be distributed (1) as at present, (2) to increase the electricity-generating capacity, (3) to supply direct heat to industry, (4) to increase gas production, (5) to produce synthetic liquid fuels for transport, particularly aviation.

Of these, (5) may be particularly important, and for the present it will be assumed that half the final energy is allocated to this with continuous production, and half to the manufacture of gas. Assuming an 80% conversion efficiency for gas^{4,5} and a similar figure for oil, the total equivalent electrical energy for each is 0.092×10^9 GJ. Since the gas contribution may be reserved, for example for on-peak domestic use, it could reduce the short-term fluctuations in the remaining demand curve, but could only reduce the annual fluctuations if gas stores (such as exhausted North Sea oil wells) or coal reserves and increased conversion plant capacity are provided; any resultant reduction in the annual fluctuations would only be possible, however, if a proportion of houses were equipped with both electrical and gas heating systems. It will therefore be assumed that long term storage is unlikely to be economic, but that a useful reduction in the peak winter loads can be achieved. The effects on the total

Table 1 Demand fluctuations (GW) and annual energy requirements

	Annual energy requirements (10^9 GJ)	Winter				Summer	
		Day	Night	Day	Night	Day	Night
Domestic:							
Space heating	0.241	49	—	12	—	—	—
Hot water	0.087	1.7	3.3	1.7	3.3	1.7	3.3
Public buildings and so on:							
Space heating	0.233	48	—	12	—	—	—
Hot water	0.038	—	2	—	2	—	2
Industry:							
Plant	0.72	32	14	32	14	32	14
Space heating	0.20	41	—	10	—	—	—
Transport	0.34	5.5	16.5	5.5	16.5	5.5	16.5
Present gas-turbine generators	0.24	36	—	7.5	—	8.7	—
Total:	2.10	213	35.8	80.7	35.8	47.9	35.8
~150-h Storage			91			42	
Annual mean					66.5		
Assuming 50% increase in coal production	1.91	187	31	79	31	47	31
~150-h Storage			82			39	
Annual mean					60.5		

The intermediate values for the spring and autumn quarters are not tabulated, but are shown in Fig. 1.

residual energy demands are shown at the bottom of Table 1. It can be seen that, while a 50% increase in coal production would make a valuable reduction in the total additional electrical energy needed (besides supplying the chemical industry), large fluctuations on all timescales still remain.

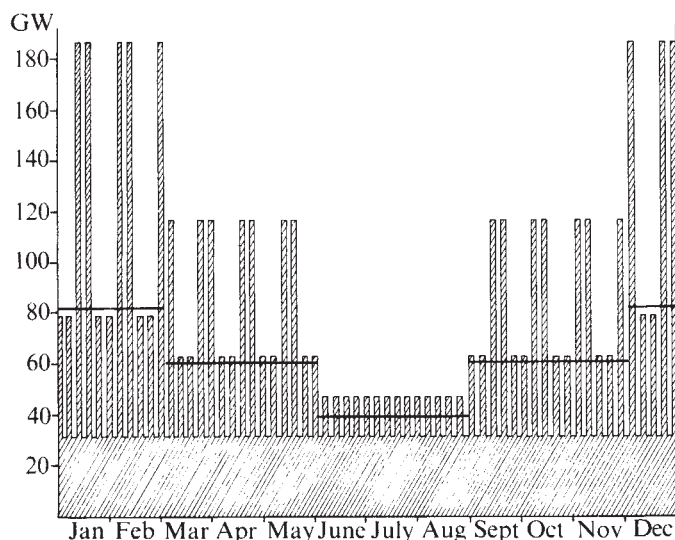
These final figures will be used in the subsequent discussion, and are shown diagrammatically in Fig. 1; it is clear that a very large installed generating capacity would be needed to meet winter peak demands—more than 3 times that needed to supply the mean energy. It is, however, evident that the inclusion of storage can make a dramatic decrease in the installed capacity needed; the provision of ~ 150 -h storage capacity would smooth out both the 24-h fluctuations and those associated with the fluctuations of outside temperature, and permit a 55% reduction in generating capacity.

Central storage systems, whether in the form of pumped hydroelectric installations (for which there are limited further economic or environmentally-acceptable sites⁴), new types of battery, compressed air, hydrogen production or flywheels, all entail a threefold conversion of the energy and are therefore of poor overall efficiency. The major cause of the fluctuations is, however, associated with the provision of low-grade space heating for domestic and other buildings, so the natural solution, both cheap and efficient, is the storage of energy as heat at the point of consumption. The extensive use of heat-pumps could further reduce the peak demand when used in conjunction with such storage, although the efficiency is lowest in cold weather, just when the demand is greatest. The storage of heat over still longer periods has been discussed, but does not seem to be economically feasible^{5,8}.

As soon as storage on the ~ 150 -h time-scale corresponding to that of weather patterns is introduced, however, the main, objection to alternative, 'renewable', sources (that is the variability of output) disappears and it becomes important to examine both the annual variations in the output from such sources, and their capital and running costs relative to those of a nuclear system. Although solar energy has an availability even less appropriate than the constant output of a conventional power station, the outputs of wind- and wave-power systems have maxima during the winter months, in line with the variations of demand.

An attempt is made here to establish the capital and operating costs of both nuclear and alternative systems

Fig. 1 Diagrammatic representation of the fluctuations in demand of the energy at present supplied by oil and natural gas; the timescale of the 24-h and 5–10-day variations has been expanded for clarity. The heavy line indicates the smoothing produced by 150-h storage.



required to supply the demand curve of Fig. 1. Fairly accurate estimates are already available for the costs of solar heating and wind-turbine generators, because both have been investigated and used in other countries for a number of years. Wave-power, tidal and other hydroelectric schemes, and other applications for solar energy will require much detailed study before reliable figures for the energy potential and the capital and running costs become available.

It is assumed that nuclear and wind-turbine generators, and solar heating panels, can all be built with useful lives of 20 years.

A nuclear system

It is difficult to establish the cost of a nuclear power system. A 1975 price for SGHWR power stations based on projections from 1973 figures has recently been given⁹ as \$1,050 per kW of installed capacity, in close agreement with a more reliable contract price (July 1976) of $\$3.2 \times 10^9$ for a 2.4 GW station for New York State¹⁰; the latter figure corresponds to a 1975 price of \$1,130 per kW. This figure does not include the capital costs of the associated fuel re-processing plant, nor of the waste-disposal facilities; details of the latter do not seem yet to have been decided¹¹. The demolition cost of a fossil-fuel station at the end of its useful life is small compared with the capital cost of a replacement, but this may not be true for the demolition of a reactor. It seems unlikely that a value to include the above points will be less than \$1,250 per kW of installed capacity, and this figure will be adopted.

The reliability of nuclear power stations has been discussed by Sorensen¹² on the basis of 1974 international data¹³ and a 1975 survey of US commercial nuclear stations has been made by the Council on Economic Priorities¹⁴. The results represent 'capacity factors' (that is the energy actually generated compared with the maximum possible) of 55% and 60% respectively; the results are worse for the larger stations. While some of the lost capacity was due to routine maintenance, some of which could be restricted to the summer months, it seems unlikely that winter capacity factors $> 70\%$ will be achieved, and this figure will be used.

The true running costs of a nuclear power system cannot be assessed until a full analysis has been made of the problems of transporting fuel, the insurance charges which might have to be covered and the cost, both in manpower and energy, of providing protection for the indefinite future, of the waste-disposal sites.

The cost of energy from UK coal, oil, nuclear and hydro power stations during 1974–75 (ref. 1), (increased by 10% to allow for inflation to mid-1975), was \$3.92, \$6.6, \$1.19 and \$0.30 per GJ respectively, and this is assumed to reflect their total running and fuel costs. In the case of fossil fuel stations, most of it is attributable to fuel costs, but in the case of nuclear stations the subdivision cannot be made until the type of reactor, the corresponding fuel processing costs, and the future effective costs of lower grade ores etc. are known. For the present it will be assumed that the nuclear fuel costs should be divided equally between fuel and running costs. Since some account must be taken of the additional costs mentioned above, a total of \$1.25 per GJ will be assumed, half being related to the installed capacity and half to the energy produced.

Various estimates have been made¹⁵ for the cost and efficiency of energy storage systems, including batteries, compressed air, hydrogen production and flywheels, and range from \$18 to \$45 per kWh; all these systems are, however, intended for reconversion to electrical power. While some central storage may be necessary, the variations in demand are dominated by the low-grade ($< 100^\circ\text{C}$) energy needed for heating buildings, for which cheaper, local, storage may be used. Such a system has recently been installed in the UK at a 1975 price of \$1.5 per kWh (J. Keable, personal communication).

Table 2 The capital and running costs (in 1975 \$) of producing 1 GJ annually with the time variations shown in Fig. 1; for all non-nuclear systems costs include storage

	Capital cost	Annual running cost	Total for 20 years
Nuclear (a) no storage	174	3.6	246
(b) 150-h storage	90	2.1	132
Solar panels	85	0.5	95
Wind	49	0.8	65
Wave	[125]	[3]	[185]

It is now possible to determine the installed capacity, and hence the capital required to provide an annual energy of 1 GJ with the time variations shown in Fig. 1, (1) without storage, (2) with storage at \$2.0 per kWh (allowing for some central storage) (Table 2). The corresponding figures for the running cost per GJ and the total of capital and running costs over the 20-year life of the station are also given; no account has been taken of inflation, since the figures are intended merely to give some idea of the relative importance of capital and operating costs for the different systems.

Alternative energy sources

The first alternative energy source I shall consider is solar energy. The mean annual variation of incident solar radiation (at Kew) is shown in Fig. 2a for (1) a horizontal collector and (2) a vertical collector facing south⁷. About half of the collected energy is due to direct sunlight and half to diffuse light. When used for providing low-temperature heat, simple collectors may be used with typical efficiencies of 40–50% (ref. 5).

This curve is not well matched to the annual variations of Fig. 1, although solar panels are already marginally cost-effective as fuel-savers for pre-heating the supply to hot-water systems⁵. The Building Research Establishment has concluded that about half of the existing houses could each be fitted with 6 m² of panel to provide 7 GJ per annum at a capital cost of ~\$330 or \$47 per GJ; the total annual energy for the number of houses which could be so equipped was estimated as 0.07×10^9 GJ (ref. 5). For new houses, where the cost of part of the roof can be subtracted, or where the design of the house as a whole can include provision for using and storing solar energy, the cost would be appreciably less; more economical designs, both of the panel and, the auxiliary components of the system, are likely to reduce the cost even for existing buildings to ~\$25 for a total annual energy supply of 1 GJ.

In order to contribute the winter values corresponding to 1 GJ having the annual distribution of Fig. 1, the area of panel would have to be increased by a factor of 4.2 (horizontal) or 2.6 (vertical south-facing). Since not all of the houses could be fitted with the latter, an intermediate value of 3 will be assumed. The capital cost of providing 1 GJ having the time-distribution of Fig. 1, must therefore be taken as \$75, although the total annual energy provided is now 0.21×10^9 GJ; the excess energy is not necessarily wasted when 150-h storage is taken into account, since at all times of the year averaging of sunny/cloudy conditions increases the useful energy abstracted, except in summer when the output may be larger than can be used for hot water heating.

Similar installations on public and industrial buildings requiring water heating could increase the total energy to ~ 0.3×10^9 GJ annually, and contribute winter powers equivalent to 0.1×10^9 GJ in the form of Fig. 1. The capital cost per GJ of the latter, including storage, is given in Table 2, together with the small running costs for which a token figure of \$0.5 per GJ is assumed. Other possible developments of solar power are discussed later.

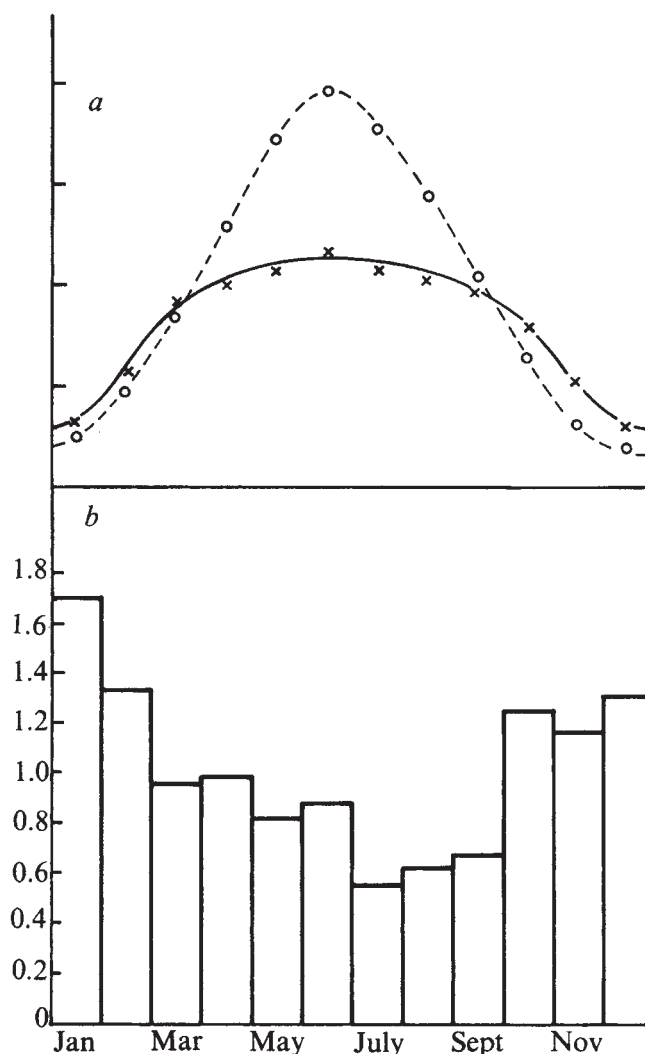
The second possibility is wind power. The annual variation

of the energy available from the wind has been discussed in a paper by Golding and Stodhart¹⁸. The most useful range of wind-speeds (containing some 80% of the total energy), is typically 5–15 m s⁻¹ in the UK and data collected at 7 meteorological stations over periods up to 10 years have been used to derive the useful energy in each month¹⁶. For the five sites situated around the coastal belt and most suitable for wind power, the average annual variation (expressed as a percentage of the mean) is shown in Fig. 2b. The data from two Midland areas with low wind speeds, and one in the Western Isles with very high wind speeds, have been excluded, although in fact the curves are similar.

A comparison of Fig. 2b and Fig. 1 shows that the annual variation of wind power is well matched to that of the demand and, since the variations of wind speed occur with the same 5–10-day periods typical of UK weather, the use of a storage system similar to that described will equally remove these, provided that a more flexible method of load-switching is introduced. Detailed records are available from four of the five sites¹⁶, and have shown that the longest periods of calm (< 4.5 m s⁻¹) vary from 18 to 209 h (the latter over a time span of 10 years); the longest period in the winter quarter with 118 h. Since these calm spells are frequently restricted geographically¹⁷, the actual variations in power available from a distributed wind-power system connected to the grid would be further reduced.

A detailed analysis of the power available annually from

Fig. 2 a, The annual variation of the energy available from solar radiation. ○, Horizontal collector; ×, vertical, south-facing collector. **b**, The annual variation of the energy available from wind power, expressed as a fraction of the mean energy.



a small domestic wind-turbine of given swept area has been made by Rayment¹⁸. Using his figures for the wind at a height of 10 m (corrected for the wind at the height of 35 m appropriate to the large turbines, 55–60 m in diameter, discussed below), the energy available is such as to provide an annual electrical output of $\sim 2.3 \text{ GJ m}^{-2}$ of swept area in the east and southern coastal regions, and 3.2 GJ m^{-2} in the western coastal regions, with values $>5 \text{ GJ m}^{-2}$ in N.W. Scotland and N. Ireland.

The vertical mixing of the air allows one to site wind-turbines so that their swept area corresponds to between 1.5×10^{-2} and 1.5×10^{-3} of the ground area (refs 19, 20, 21, 22); taking a value of 4×10^{-3} and a mean density of 2.8 GJ m^{-2} , the annual electrical energy would be $\sim 11,200 \text{ GJ km}^{-2}$. These coastal areas have a total area of about $1.5 \times 10^5 \text{ km}^2$.

The whole area could not, of course, be used, though the impact of such wind-turbines, e.g. on agriculture, is minimal. (If 50–60 m diameter turbines ($\sim 1 \text{ MW}$) were used, the structure would be comparable with a 275-kV grid-pylon and their separation would be $\sim 1 \text{ km}$. Since local connections would be at medium-voltage, their visual impact would, in fact, be little different from a nuclear system in which most of the power stations would have to be sited on the coast; the number of (275-kV) pylons within the coastal belt for the latter system would be about two-thirds that for the wind system).

In some areas, as with grid lines, there would be strong environmental objections although, unlike those which apply to a nuclear system, they would be limited to their visual impact. For the present it will be assumed that only one-third of the coastal area considered could be used, with an annual output of $0.56 \times 10^9 \text{ GJ}$. It has been suggested¹⁹ that wind-turbines could also be built in shallow and accessible off-shore areas, such as the Wash and Thames Estuary; while the capital costs of building the turbine and of transmission ashore are greater than those for a land installation, the greater wind speeds would provide an annual energy ~ 5 to 10 GJ m^{-2} instead of the value of 2.8 GJ m^{-2} assumed for the coastal area; the capital costs per GJ may therefore be no greater than for the land installations considered. It seems probable that a further $0.4 \times 10^9 \text{ GJ}$ could readily be provided in this way. The total of $0.96 \times 10^9 \text{ GJ}$ having the annual variation of Fig. 2b would provide an average power of $\sim 43 \text{ GW}$ during the winter quarter, and $\sim 20 \text{ GW}$ during the summer quarter.

Whereas an increase of scale introduces significant economies in conventional power stations, it is not yet clear that this is the case for wind-turbines much greater than 1 MW. The following analysis is therefore based on turbines up to 1 MW for which considerable information on design and cost is now available; for example several installations with capacities up to 200 kW were built in Denmark, and 800-kW and 1-MW turbines were built in France, and were used to provide power for their respective grid-systems during 1955–65^{15, 23, 24}. The Danish report²³ concluded that by 1975 wind-turbines of this size were cost effective relative to oil-fired power stations even as fuel-savers. The capital cost of their 1957 experimental 200-kW design installed at Gedser, when converted to 1975 prices, was $\$280$ per m^2 of swept area for production on a limited scale. A more modern design would have $\sim 40\%$ greater efficiency²⁵, so this figure, with UK wind speeds, would correspond to about $\$700$ per kW of installed capacity. It is, moreover, likely that a modern wind turbine would be cheaper, largely owing to the development of reliable solid-state devices, which allow the use of a more efficient turbine blade running at fixed pitch (thus removing one of the most serious stress problems) and driving a variable-speed alternator by way of a simple fixed-ratio gearbox, thus avoiding both the complexity of variable-ratio gearboxes and the need for synchronisation with the grid. The output of

such an alternator can easily be varied to present the optimum torque-speed relationship, and would probably be converted to fixed voltage d.c. and transmitted to a sub-station collecting the output from a number of installations for inversion and supply to the grid. A cost significantly below the figure of $\$700$ per kW is therefore to be expected.

The analysis of the Mitre Corporation¹⁵ gives a figure of $\sim \$500$ per kW for a 100-kW installation and $\$300$ per kW for a 1-MW installation—again for a rated wind speed (12 m s^{-1}) appropriate for the UK. Recent design studies²⁵ for 1.5-MW turbines, for lower rated speeds (10.7 m s^{-1}) but rather higher mean wind speeds (8 m s^{-1}) by the General Electric and Kaman Companies in the USA gave costs of $\$449$ and $\$481$ per kW of installed capacity; these use a synchronous alternator which, as discussed above, is not likely to be the most economical system, except for use over a narrow range of operating wind speed. A recent design study by Wesco (G. W-W. Pontin, personal communication) of Redhill for a 46-m diameter turbine and for the assumed UK wind pattern gives a figure of $\$275$ per kW.

All the above systems are based on conventional horizontal-axis turbines; it is possible that more economical designs using a vertical axis will be developed^{19, 26}.

If the local transmission and substation are included, it is therefore reasonable to adopt a figure of $\sim \$500$ per kW of installed capacity. The remaining transmission costs, as already discussed, are not likely to exceed those of a nuclear system.

From statistics of the wind and an assumption of 90% mechanical serviceability, it is now possible to compute the installed capacity, and hence capital cost, of a wind system needed to provide each GJ with the annual variation of Fig. 1. This figure when added to the storage cost, is included in Table 2.

The Danish report concluded that the running cost of a wind system would be $\$0.93$ per GJ, although this figure is probably an overestimate in the light of the improvements and simplification of design discussed above; the corresponding figure from the Wesco data is $\$0.4$ per GJ. A value of $\$0.8$ will be used.

The third possibility is wave power. A useful survey of the energy available from wave power has been made by Glendenning and Count²⁷ who give, for the West coast of the British Isles, average annual powers of 50–100 kW per m of collector. The medium-term fluctuations in power are very similar to those for overland wind, with a marked variability on a timescale of 5–10 days, while the average available power in winter is again about $2\frac{1}{2}$ times that in summer. Since much of the energy originates in waves generated at distances $>1,000 \text{ km}$, the 5–10-day fluctuations will not be correlated with those of overland winds, so that a combination of wave and wind power will show smaller fluctuations than either alone.

The cost has been estimated as $\$600$ – $\$1,200$ per kW of installed capacity²⁷ but this estimate must be regarded as provisional; work supported by the Department of Energy at the University of Edinburgh, the National Engineering Laboratory, and Wavepower Ltd should lead to a more accurate figure after study of the practical problems of mooring such devices in deep water, their survival during storms and the restrictions and dangers to shipping; since some 300 km of collector would be necessary to equal the annual power output of the wind system considered, this latter point is clearly important.

The problems of bringing the power ashore are also much more severe than those involved in either nuclear or wind systems, whether an electrical cable, piped-hydrogen or any other method is used; for the present a capital cost of $\$1,500$ per kW will be assumed, with a serviceability of 80%. The running costs and maintenance of the collector structure will also involve a considerably greater cost per GJ than for a wind system.

Implications for bridging the energy gap

The large fluctuations in the energy demand at present supplied by oil and gas require a generating capacity, for an electrical alternative, some 3 times that needed to supply the mean energy. It does not seem feasible to build the ~ 266 GW of nuclear generating capacity required before the end of the century. If storage, of ~ 150 h duration, is included into the system to smooth out the diurnal and weather fluctuations, the installed capacity could be reduced to ~ 117 GW; the cost of the storage, most of which would be at the point of consumption, would be much smaller than that of the generating capacity saved.

As soon as such storage is introduced, however, a number of the alternative sources of energy can be compared, on an equal basis, with a nuclear system. Table 2 shows that wind energy is the cheapest of all the 'renewable' sources developed so far, and that a system based largely on wind turbines is very much cheaper, both in capital and running costs than a nuclear system; since the figures are based on experience with actual machines comparable in size with those envisaged, any uncertainties in the costs are unlikely to alter this conclusion.

The energy curve produced by the example wind system described is given in Fig. 3, which shows that it would supply half of the demand throughout the year.

The installation of solar panels also provides cheap energy although its poor fit to the annual demand curve increases the effective cost significantly above that of a wind system. It would be useful mainly for hot-water supply, and the installation discussed would contribute the energy curve also shown in Fig. 3. It can be seen that these two sources would together provide a major part of the total demand. The remainder might be obtained either by extending the wind system or by wave power or new types of hydro-electric system²⁸. For the particular case of fuels for transport, plant photosynthesis may become cost-effective. (For example, sugar-beet converts 4.3% of solar energy while methane production from fermentation of sewage, straw and crops such as *Spartina*, which grows on salt-marshes, would not require the use of land suitable for food production⁷).

The actual introduction of the wind system (and solar panels) could be straightforward and quick. An alternative off-peak load-switching system would be needed in place of the rigid time-switches at present in use. The replacement



Fig. 4 An experimental 3-m (2-kW) wind-turbine at Cambridge used for the design of efficient blade profiles and methods of load matching and for the investigation of the non-steady loads due to gusts.

of these clocks with a (probably cheaper) detector of a high-frequency carrier transmitted via the mains from the sub-station (as has been used for street lighting and other switching operations), would allow the load to be adjusted to meet the energy available whatever its time or geographical variations; such a system would also have the advantage of allowing better use of existing generating capacity (for example, weekend surplus, avoidance of high peaks at the start of the off-peak period when thermostat switches are all closed).

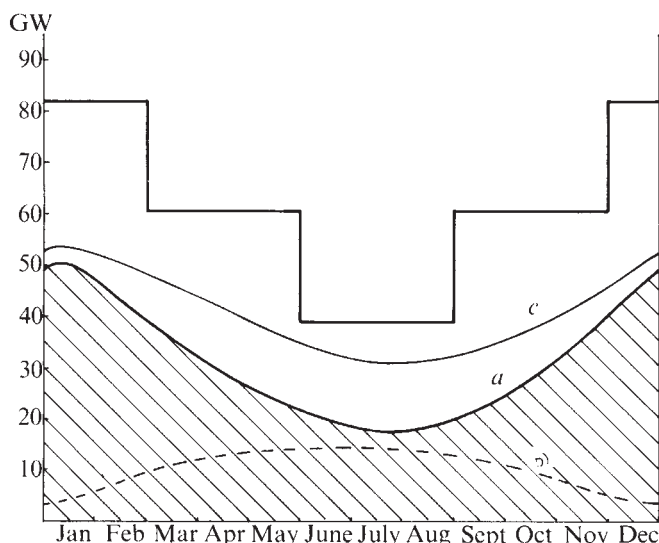
As with all alternative systems a single prototype unit of small cost coupled with site surveys allows a reliable estimate of the cost and performance of a complete system, in marked contrast to the expense of building a prototype reactor. In the development both of a prototype wind-turbine, and in subsequent production, about 60% of the cost is accounted for by the tower, gearbox and alternator¹⁹, all of which are virtually 'off-the-shelf' items. The blade design seems to be a problem well within the capabilities of the British aviation industry on a timescale of 6–12 months and a production prototype could be operating within 18 months with the expenditure of perhaps \$1–2 million (the first full-scale Hovercraft flew 8 months after design started).

The subsequent production could build up over a 5-year period to 2,000 units a year (representing, in both cost and tonnage of steel, 10–15% of that of the UK motor industry) and divided approximately equally between the structural steel, the electrical and the aircraft industries. By 1985 wind energy could be providing $>10^7$ GJ annually and the energy consumed in the manufacture of each unit² would be repaid by it in 6–12 months.

In contrast, unless large scale production of an existing reactor design were adopted, the introduction of a nuclear system is inevitably slow. In the case of a new design, such as the proposed FBR (for which there may in any case, be a limitation of $\sim 3\%$ per year in the increase of the system output due to limitations of fuel supply^{11,29}) it would be necessary to build and test, on a timescale of 8–10 years, a prototype before construction of further stations can proceed.

If such a prototype (say, 2–3 GW) were in operation by 1986–87, and a period of perhaps 3–5 years were necessary to repay the energy involved in its construction, fuel extraction and processing and so on, it would only begin to supply useful energy (as Fig. 1) by about 1991, at an annual rate of $\sim 2 \times 10^7$ GJ; the wind system discussed, with a production

Fig. 3 The total energy (in relation to the smoothed demand of Fig. 1) provided (a) by the wind system described (shaded), (b) by the installation of solar panels as described, and (c) by both combined.



of 2,000 units p.a. would, by then, be producing an annual output of about 2×10^8 GJ.

Not only does wind energy seem to be the cheapest way of reducing the energy gap, it is the only way of making a significant contribution within the next 12–15 years, and so allow the preservation of some of the North Sea oil for the future.

Although the discussion has been limited to the specific problem of UK energy supplies, the conclusion may have even greater application in other parts of the world, most of which have either adequate wind or solar resources to support the needs of a lower density of population; the straightforward technology requires only simple maintenance and none of the security problems of nuclear stations. The possibilities of very large export markets should add further incentive to develop the first production version of a 1 MW wind-turbine.

I thank Dr A. J. P. Winter for assistance in preparing this paper.

¹ United Kingdom Energy Statistics 1973–75. (HMSO, London, 1976).

² Increased Costs of Energy; Implications for U.K. Industry. (NEDO, HMSO, London, 1974).

³ Energy Conservation in the U.K. (NEDO, HMSO, London, 1975).

⁴ Energy in the 1980s. *Phil. Trans. R. Soc. A* 276, 405 (1974).

⁵ Energy Conservation: a Study of Energy Consumption in Buildings and Possible Means of Saving Energy in Housing, CP 56/75. (Building Research Establishment, 1975).

⁶ Lloyd, S. A. & Weinberg, F. J. *Nature* 257, 367 (1975).

⁷ Solar Energy, a U.K. Assessment. (UK Section of ISES, London, 1976).

⁸ Yellot, J. I. *ASHRAE Journal*, December 31 (1973).

⁹ Wilson, E. M. *Symposium on the Potential for Power from Currently Unconventional Energy Sources* (Institute of Fuel, Southampton, in the press).

¹⁰ *New York Times* 23 July (1976).

¹¹ Royal Commission on Environmental Pollution, Sixth Report; Nuclear Power and the Environment (HMSO, London, 1976).

¹² Sorensen, B. *Science* 194, 135 (1976).

¹³ Operating Experience with Nuclear Power Stations in Member States in 1974. (International Atomic Energy Authority, Vienna, 1975).

¹⁴ Power Plant Performance: Nuclear and Coal Capacity Factors and Economics. (Council on Economic Priorities, USA, 1976).

¹⁵ Eldridge, F. R. *Wind Machines* (Report to the NSF, Mitre Corporation, USA, 1975).

¹⁶ Golding, E. W. & Stodhart, A. H. *The Potentialities of Wind-power for Electricity Generation*, Report No. W/T16 (BEIRA (now ERA), 1949).

¹⁷ Molly, J. *Proc. Int. Symp. Wind Energy Systems*, Cambridge, UK (BHRA Fluid Engineering, Cranfield, in the press).

¹⁸ Rayment, R. CP 59/76. (Building Research Establishment, 1976).

¹⁹ Musgrove, P. *New Scientist* 72, 596 (1976).

²⁰ Coty, U. & Dubey, M. *2nd Ann. Energy Symp.* (Los Angeles Council of Engineers and Scientists, Los Angeles, 1976).

²¹ Ljungstrom, O. *Proc. Int. Symp. Wind Energy Systems*, Cambridge, UK (BHRA Fluid Engineering, Cranfield, in the press).

²² Raily, J. W. *New Scientist* 73, 231 (1977).

²³ Wind-power Rep. Danish Academy Technical Sciences (1975).

²⁴ Wind Workshop 2 *Proc. of the ERDA–N.S.F. Conf. Washington*, (Mitre Corporation, 1975).

²⁵ Thomas, R. L. *Proc. of the Vertical-axis Wind-turbine Workshop* (E.R.D.A., Albuquerque, 1976).

²⁶ Templin, R. J. *National Research Council of Canada. Report LTR-LA-160* (1974).

²⁷ Glendenning, I. & Count, B. M. *Proc. Symp. Renewable Sources of Energy R. Soc. Arts*, p.50 (1977).

²⁸ U.K. Prov. Patent Spec 10006/77.

²⁹ Merrick, D. *Nature*, 264, 596 (1976).

articles

A chaotic cosmology

John D. Barrow

Department of Astrophysics, Oxford University, Oxford, UK

An anisotropic, inhomogeneous cosmological model is proposed in which the inhomogeneity is generated by shear fluctuations. This is a sufficient condition for dissipative heating by collisional neutrinos to explain the present large heat content of the universe, $S_{\text{b}0} \sim 10^8$, together with its isotropy and comparative homogeneity on large scales when the photons were last scattered. The model does not require the chaotic motions to be arbitrarily truncated on large scales and isotropises early enough with high entropy to ensure the synthesis of light elements with the observed abundancies. A population of black holes which arises in a natural way can also provide the necessary ingredients for a theory of galaxy formation and morphology. The 10^{15} -g black holes, predicted by some authors, are not necessarily expected to be a feature of chaotic cosmologies.

MISNER¹ first pointed out a fundamental problem which measurements of the microwave background isotropy had created². He drew the attention of cosmologists from seeking an explanation for the presence of local enhancements over the background density on galactic scales in terms of random fluctuations to asking why the background universe should display a high degree of isotropy and homogeneity at all. This approach to cosmology was termed 'chaotic cosmology' and in its most general form seeks to show that the observed structure of the universe is demanded rather than merely permitted by the physical processes acting within it. It was thus hoped that

general relativity would predict that the universe is unique in theory as well as in practice, irrespective of its initial state.

Early attempts by Misner to resolve, within the framework of classical relativity, the problems of the contingency of isotropy and the apparent non-intersection of past light cones, together with the fundamental difficulties they encountered at a classical level are well known (see ref. 3). It is still hoped that a complete resolution of the problem will come by appealing to quantum phenomena in the first 10^{-43} s of the Universe's life, but as this route is at present still beset by many conceptual and technical problems⁴ it is interesting to investigate what may be achieved by known classical effects and it is at this level that we address the problem.

The most important unexplained properties of the universe for which cosmologists seek an explanation are probably: (1) The large scale homogeneity and isotropy as indicated directly by the absence of structure in the intensity of the microwave radiation on large angular scales $(\Delta T)/T \leq 3 \times 10^{-3}$ (ref. 2), and indirectly by the observed cosmic mass fractions of helium ~ 0.3 and deuterium $\sim 2 \times 10^{-5}$ predicted by the open homogeneous, isotropic Robertson–Walker (RW) cosmological model⁵. (2) The large heat content of the background radiation, dimensionlessly characterised by the number of black-body photons per baryon, $S_0 = (4a/3k)(T_{\gamma 0}^3/n_0) \sim 10^8 \gg 1$ (where a is the radiation constant, k the Boltzmann constant, $T_{\gamma 0} = 2.7$ K, the present radiation temperature and $n_0 \sim 10^{-5} \text{ cm}^{-3}$ the present baryon number density). This is unusually large for a dimensionless physical quantity, but far smaller than the other large cosmic numbers $\sim 10^{40}$ (ref. 6).

The existence of such a uniform physical system in this high entropy state is highly suggestive of there having been a con-

siderable amount of dissipation in the past, although the finite value of the entropy indicates that there could not have existed an arbitrarily chaotic initial state³². (3) The origin of protogalactic inhomogeneities with amplitudes sufficiently large to generate distinct structures on scales $\sim 10^{11}$ – $10^{15}M_{\odot}$ by the present era, yet sufficiently small to avoid a proliferation of compact objects in the same mass range. (4) The absence of developed structure on scales larger than $\sim 10^{15}M_{\odot}$ and the apparent absence of a preferred scale of galaxy clustering over the range 10^{12} – $10^{15}M_{\odot}$ (ref. 7). In conjunction with (3) this presents the problem of explaining the observed spectrum of inhomogeneities at recombination.

The aim of this paper is to outline a chaotic cosmological model in which the above features of the universe might be explained by a hypothesis concerning the type of non-uniformity present at the singularity. There have been many attempts to explain the above problems (1)–(4) in isolation; for example the work of Misner and Matzner⁸ on (1); Rees⁹, Zeldovich¹⁰ and Liang¹¹ on (2). Many workers have discussed aspects (3) and (4) (ref. 12).

The model described here owes much to the chaotic cosmology conceived by Rees⁹, combining other features to improve and extend it.

The Rees cosmology

In brief, the Rees cosmology consists initially of an approximately isotropic distribution of baryons with large amplitude inhomogeneities, $\delta\rho/\rho \sim 1$, on each scale as it enters the horizon. The radiation content of the universe is then generated exothermically by the nonlinear dissipation of the acoustic waves as their scale is encompassed by the Jeans length. For a state of maximal chaos the energy budget is given by $aT^4 \sim \rho_b c^2$, which gives an entropy production–redshift relation, $S = 10^{13}(1+Z)^{-1}$. The shocking of acoustic fluctuations can thus generate the radiation content of the universe by a redshift, $1+Z_f \sim 10^6$. Rees points out that the expected small scale matter clumping on scales less than the photon diffusion length enhances the free–free production rate and enables the created radiation to be adequately thermalised at this redshift. Interestingly, the clumping of matter on scales larger than the photon diffusion length would, in complete contrast, make thermalisation much more difficult than in a smooth model of the same average matter density. This effect could provide an interesting constraint on the existence of lagging cores in the early Universe as envisaged by Harrison¹³.

To prevent the production of more entropy than is observed in the microwave background one must truncate the chaos on the horizon scale corresponding to $1+Z_f \sim 10^6$ encompassing $\sim 10^{17}M_{\odot}$, which is close to the mass of the largest observed structures in the universe. This scheme also explains the approximate coincidence of the epoch of recombination and the onset of matter domination, $Z_{eq} = 10\Omega Z_{rec}$ where Ω is the ratio of the present baryon density to the closure matter density.

Attractive as this conception is, it does however possess a number of undesirable features.

(1) The presence of inhomogeneities having $\delta\rho/\rho \sim 1$ on horizon sized scales with their associated lightlike peculiar velocities is not consistent with the assumption of global isotropy. As a consequence the expansion timescale will accordingly be shortened to correspond with a locally anisotropic (say, Kasner-like) structure for $Z \geq 10^6$. Some of the implications of this situation are mentioned below.

(2) The final entropy, $S_b \sim 10^8$, is principally generated by the dissipation of the largest chaotic mass scales. For shock dissipation the entropy production obeys $S = 3 \times 10^{13} T^{-1}$. This means that during the period $Z \sim 10^9$ – 10^{10} , conducive to light element synthesis, the specific entropy $\sim 10^4$ is too low to produce the observed deuterium abundance $\sim 10^{-5}$, under-producing it by a factor of $\sim 10^8$ and also producing a dis-

cordant helium fraction of 40% in excess of the observed 30%. If one also takes into account the accompanying effects of anisotropy the discord becomes even worse, three anomalous situations being possible according to the degree of anisotropy at the time of nucleosynthesis, $T \sim 10^9$ – 10^{10} K. First, a helium abundance of ~ 60 – 100% in a state of high anisotropy and thermal equilibrium. Second, indeterminate abundances of all light elements if the expansion time scale is less than the weak interaction time but greater than the slowest nuclear reaction time scale. And third, zero abundances of all light elements heavier than hydrogen when the expansion time is less than both weak and nuclear reaction times.

The latter two possibilities can arise in models in which matter motion at near light-like velocities leads to a Lorentz contraction of the expansion time scale in the rest frame of the moving matter^{14,33}.

(3) There is the necessity for an arbitrary cut-off to the chaotic motions on scales larger than $\sim 10^{17}M_{\odot}$. No physical mechanism is suggested which could bring about such a cut-off. The only possible mechanism would seem to be a photon ‘viscosity’ due to the long mean free paths of decoupling photons. The scattering of photons by non-relativistic electrons would give rise to a constant viscosity $\eta = m_e c p_e (kT \sigma_T n_e)^{-1}$ apparently capable in theory of damping homogeneous shear of infinite wavelength. The recombination of the plasma takes place too rapidly for there to be a viscous period of any significance, however, and any extension of the viscous epoch due to the dissipative heating would seem to be prohibited by the Planckian form of the observed spectrum¹⁵.

(4) The degree of large scale anisotropy compatible with the required inhomogeneity at $Z \sim 10^6$ would almost certainly do violence to the observed constraints $(\Delta T)/T \leq 3 \times 10^{-3}$, on the microwave anisotropy at last scattering unless there occurred sufficient reheating of the cosmic plasma at late times to defer its last scattering until $Z \sim 10$. There are, however, strong limits to the amount of anisotropy energy which can be ‘buried’ in the microwave photons by dissipative processes at late times since distinctive spectral distortions can be expected to arise¹⁵. B. J. T. Jones has pointed out that these distortions cannot yet be predicted straightforwardly as the Kompaneets’s equation describing the evolution of the general photon spectrum is of a different character in a non-uniform environment¹⁶. One would not expect anisotropy to alleviate this distortion in the way suggested by Rees³¹ since the anisotropy must be damped out before the photons are thermalised.

Prolific black hole formation on all mass scales up to $\sim 10^{17}M_{\odot}$ might also be expected and would be in glaring contradiction to observation¹⁷.

(5) The results of Peebles¹⁸ and E. T. P. Liang (preprint, 1976) indicate that difficulties are encountered if one attempts to explain the observed structures on globular cluster and galactic scales as a residue of nonlinear damping. Shocking is highly efficient as an eradicator of large amplitude, short wavelength fluctuations, damping out the adiabatic fluctuations on relevant scales $\sim 5 \times 10^{11} \Omega^{-1} M_{\odot}$ completely and one must appeal to fragmentation of large $\geq 10^{14}M_{\odot}$ clumps or isothermal perturbations to describe the presence of galaxies. In fact the damping is probably much more efficient and even larger scales removed than those predicted by Liang since his calculation based upon a simple wave picture in which the self-interaction of single sound waves is adequately described but their mutual interaction is not. A more complete picture of ‘acoustic turbulence’ including these effects might give rise to more efficient dissipation over larger scales than the simple wave picture.

A new model

We describe here an initially anisotropic and inhomogeneous cosmology which generates the observed microwave radiation by the dissipation of chaotic motions but in which the chaos is

damped out on all scales at early times. In this way one enables light element synthesis to proceed as in the standard RW models and the radiation background to become isotropic with Planckian spectrum without having to invoke small scale clumping or re-ionization at late times.

We shall take the singularity structure to be defined by an anisotropic and inhomogeneous zero entropy state filled with pressureless dust. This is locally equivalent to a Kasner universe with large amplitude, $\delta\rho/\rho \sim 1$ matter perturbations on all scales when they enter the horizon. The work of Belinskii *et al.*¹⁹ indicates that such a representation is sufficiently general to constitute an open neighbourhood in the superspace of all cosmological solutions to the Einstein equations. It is also a self-consistent picture since the presence of global spatial inhomogeneity implies anisotropy in the expansion dynamics. We now hypothesise that the inhomogeneity be limited to that class which is generated by equilibrium shear fluctuations, the meaning of and reason for this restriction will be explained below.

The very early evolution of this model is envisaged as proceeding in similar manner to the Rees cosmology, entropy being generated by relativistic shock damping of the large amplitude acoustic waves, leading to a growth of entropy according to the relation $S = 3 \times 10^{13} T^{-1}$, as above. This process proceeds efficiently since the relatively soft equation of state guarantees that the local sound speed is much less than the peculiar velocities induced by the inhomogeneity. More rigorously, solution of the fluid dynamical equations for test motion in the anisotropic background shows that the velocity field satisfies

$$U_k = f[(x_k \pm ((ct)^{-p_k})/1 - p_k)]$$

where U_k is the fluid three velocity, f an arbitrary differentiable scalar function and p_k are the usual Kasner indices satisfying $-\frac{1}{2} \leq p_1 \leq 0 \leq p_2 \leq \frac{2}{3} \leq p_3 \leq 1$; flow singularities thus develop.

This process will continue to produce entropy until the temperature falls to $\sim 10^{11}$ K when the influence of weakly interacting particles^{8,20} becomes crucial to a discussion of the dynamics. It is at this temperature that the viscous action of decoupling electron and muon neutrinos can lead to the damping of the expansion anisotropy. We consider the equilibrium neutrino case, with the neutrino-electron scattering time shorter than the expansion time, $t_{ve} < t_{exp}$, in which the corresponding coefficient of viscosity is derived from the Feynmann Gell-Mann theory of weak interactions, (20)*, with $\eta \sim \rho t_{ve} \sim T^{-1}$. (The existence of neutral currents make a negligible difference to the result.) We describe the anisotropy by the ratio of the anisotropy energy density, to that of the thermal energy, defining

$$A^2 \equiv \rho_B/\rho_V = \sigma^2(NaT^4)^{-1}$$

where σ is the kinematic shear and N the weighting function for the particle species. It can be shown that the equilibrium neutrino assumption restricts the allowed anisotropy to $A \sim 10^6$. (This is demonstrated by comparing the ratio of the neutrino-electron collision time, $t_{el} \sim 10^{59} T^{-5}$ to the expansion time $\sim (G\sigma^2)^{-1}$ at the dissipation epoch when $T \sim 5 \times 10^{10}$ K). By calculations analogous to those of Misner²⁵, it can be shown that in this case the viscous action of the neutrinos is always effective in reducing the anisotropy energy from its maximum allowed equilibrium value, $A \sim 10^6$, to $A < 1$ within an expansion time and the growth of entropy is $S/S_i \sim 10^5$, where S_i is the value of the specific entropy before the neutrino viscosity became effective. This is precisely the amplification factor necessary to augment the shock generated entropy, $S_i \sim 10^3$, to the required value, $S_0 \sim 10^8$ for the products of nucleosynthesis at $T \sim 10^9$ K to be in agreement with observation. (For a general temperature dependent viscosity $\eta \sim T^n$, we find

$$S/S_i \sim 5 \times 10^{6(1-n)/(2-n)} / T_i^{(n^2-1)/(2-n)}$$

where T_i is the adiabatic temperature at which viscous dissipation commences, so $\rho_{vi} \equiv NaT_i$. These results can be obtained from those of Matzner²⁴.) Some vorticity will also be created in the shock fronts and by the viscous dissipation itself²¹ but this is created only on $\lambda < vt$ scales where photon viscosity will subsequently damp it out and it thus has questionable relevance to turbulent models of galaxy formation.

Now we also notice that since the inhomogeneity is generated by shear fluctuations it will also be damped out by the neutrino viscosity, even on scales larger than the horizon. No acausal process is implied here, for as the infinite wavelength homogeneous shear is damped so is its global fluctuation. More formally we know that a space which is locally isotropic everywhere must be spatially homogeneous and here it is the action of the constraint equations of elliptic character that force the space into a homogeneous state. Such an effect has been found in the case of small inhomogeneity by Liang²² but was not interpreted physically and it is in fact also apparent in the results of Stormer²³. The smoothing process can also be viewed in terms of the initial value problem. The evolution equations are hyperbolic but the initial conditions are set so that non-communicating domains share some common features. In particular, fluctuations on scales smaller than the horizon are everywhere damped and hence appear damped on scales greater than the horizon (M. A. H. MacCallum, personal communication).

Thus the viscous mechanism is also effective in removing the inhomogeneity on all scales by the time the temperature has fallen to $T \sim 10^9$ K and on large $\lambda > ct$ scales the universe is then effectively the standard Robertson-Walker model—the structure on smaller scales is discussed below. The model can then generate the observed helium and deuterium abundances in the usual way with the proviso that the entropy dependence of the latter is modified by the effects of its small scale fluctuation.

Finally, we turn to the problem of the possible origin of suitable protogalactic structures. Clearly we could make the *ad hoc* assumption that there exist growing curvature fluctuations with $\delta\rho \sim 10^{-4}\rho$ on all scales when they enter the horizon as required in the theories of Zeldovich *et al.*²⁶ and Gott and Rees²⁷, which would escape any viscous dissipation on scales greater than the horizon. Such a state of affairs is unconvincing, however, and a second possibility arising naturally in the model is more interesting.

Meszaros²⁴ has proposed a theory of galaxy formation based on the hypothesis that a fraction of the matter in the early Universe collapsed to form a population of low mass black holes, the subsequent fluctuation in their number density constituting a non-decaying isothermal density perturbation. The absence of electromagnetic coupling between the black holes and the background radiation field ensures that they behave as entropy perturbations experiencing no damping by radiative pressure. Their amplitudes on scales smaller than the horizon remain frozen until the radiation ceases to dominate the expansion dynamics, after which linear growth proceeds in the usual way. This scenario is shown to lead to a consistent theory of galaxy formation and morphology at later times. In Meszaros' work the origin, number density and mass of the black holes are all arbitrary parameters chosen as initial conditions to guarantee the right amplitudes at recombination. In the model described above, however, such a black hole population arises naturally during the viscous damping period while the model is undergoing isotropisation. The shocking of inhomogeneities when they enter the Jeans length will not be completely efficient and when the sound speed is c_s we would expect some inhomogeneities of amplitude δ to collapse into black holes with a probability $\sim \delta \exp(-\frac{1}{2}c_s^2\delta^{-2})$ (ref. 17), if their amplitudes are appreciable on scales intermediate between the Jeans and horizon masses. At first sight it might seem from the results of Carr that black holes will form prolifically on all scales below $\sim 1M_\odot$ leading to $\rho_{bh} \sim \rho_{dust}$ in this mass range and thus causing the Hawking evaporation¹⁷, at the present epoch of

those with $M \sim 10^{15}$ g to violate the limits on the observed γ -ray flux in the 100 MeV range which requires ρ_{bh} (10^{15} g) $< 10^{-6} \rho_{dust}$. We claim, however, that black-hole collapse will only occur on horizon sized scales during the isotropisation and homogenisation process itself when $A \leq 1$. This will produce a population of small black holes with mass $\sim 1 - 10^4 M_\odot$ and their density must satisfy $\rho_{bh} \leq \rho_{dust}$. The reason for this restriction in the mass range is that the calculations of Carr model the inhomogeneity as a curvature fluctuation by joining different RW models together. The space-time is thus allowed global isotropy although here global anisotropy is required for self-consistency. Now, whereas in an isotropic universe the collapse time for the matter $\tau_c \sim \rho_\gamma^{-1/2}$ equals the expansion time τ_{ex} and collapse can readily occur, this will not be the case in an anisotropic background where the expansion rate is

$$\left(\frac{\rho_B}{\rho_\gamma}\right)^{1/2}$$

times faster.

Thus, as the collapse time remains $\sim \rho_\gamma^{-1/2}$, anisotropic dynamics guarantee $\tau_c \gg \tau_{ex}$ and there is insufficient time for the black holes to form. A more physical explanation is that the dominant anisotropy pressure determines the Jean's mass, making it equal to the horizon size. Since the gravitational collapse takes place on scales larger than the Jean's mass, overdense regions collapse while still larger than the horizon size to form separate universes¹⁷, and not black holes in our Universe. (This problem is now being investigated in detail with B. J. Carr.) This argument indicates that one might not expect to observe the Hawking effect today even if the early Universe were highly chaotic, unless 10^{15} -g black holes were present in the universe *ab initio*.

By generalising the calculations of ref. 24 it can be shown that if the number fluctuations in the black hole distribution satisfy

$$\frac{\delta N}{N} = N^{-n}$$

the perturbations will form discrete bound systems of mass M at a redshift

$$1 + z_b \sim 4 \times 10^{6-11} n \Omega^{3/4} t_i^{-1} \left(\frac{b}{1 M_\odot}\right)^n \left(\frac{M}{10^{11} M_\odot}\right)^{2/3-n} \quad (1)$$

where b is the initial mass of the hole and t_i the formation time. But, we know $b(t_i)$ in the isotropising model from the expression for the horizon size³⁵, $b \sim 10^4 t_i M_\odot$.

For an initially Poisson fluctuation, $n = \frac{1}{2}$, there is a minimum of parameters in equation (1) and the largest scales condense out first; where for partially correlated fluctuations, $n = 2/3$, all scales condense simultaneously. In both cases we can have $1 + z_b < 10$ for galaxy and cluster masses. This scheme of protogalaxy formation has a number of interesting points in its favour since the perturbations are not damped by photon diffusion in the radiation era, a process which smooths the interesting $\leq 10^{13} M_\odot$ scales from the spectrum of adiabatic fluctuations. Also, from equation (1) we see that there is a continuous scale of galaxy clustering and the existence of the galactic scale as a lower bound arises because gas dynamic processes are able to destroy the clustering on smaller scales. It is supposed that these non-gravitational effects imprint the characteristic scales and structure on galaxies rather than the precise spectral form of the density inhomogeneities³⁰.

There seems to be no observational evidence against a large number of small black holes having, say, the same mass spectrum and distribution as stars²⁴. The model might be open to observational test by the effect pointed out by Dahlbacka *et al.*²⁸. If gas which falls into a black hole cannot find a more efficient mechanism than electron ion collisions to maintain $T_e = T_{ion}$, (for example through turbulent or magnetic effects),

then a feature could be expected in the γ -ray background in the 10–20-MeV range.

Conclusions

I have tried to show that a partially chaotic model of the early universe can provide an explanation for the entropy and uniformity observed today together with a viable scheme of protogalaxy formation and clustering. The model has the following input. (1) The inhomogeneity is restricted to that class generated by shear fluctuations on all scales at the singularity. And (2), There is neutrino equilibrium, such that $t_{\nu e} < t_e$ which requires that the anisotropy, as measured by

$$A \equiv \left(\frac{\rho_B}{\rho_\gamma}\right)^{1/2}$$

is bounded above.

The restrictions (1) and (2) in conjunction describe 'equilibrium shear fluctuations'. Both these restrictions have phenomenological motivation. First, we know that the isotropy and finite entropy observed today limit the initial degree of chaos to a finite level³², and second, if there did not exist thermal equilibrium of the neutrino mediated weak interactions when $T \sim 10^{10}$ K then the synthesised abundances of helium and deuterium would be either indeterminate or non-existent¹⁴. The degree of generality admitted by the assumption (1) is difficult to assess precisely since no good criteria exist to determine physical cardinality. But as the influence of shear fluctuations becomes infinite as the singularity is approached whilst that of curvature fluctuations becomes negligible, the restriction (1) may in some ways be natural for a chaotic model.

Finally, two important and as yet undiscussed problems still loom large on the horizon of chaotic cosmology at a classical and a quantum level. We could ask why the discussion should be limited to baryon symmetric models of the initial state or indeed is the whole concept of a zero entropy (non-thermal) yet chaotic initial state a contradiction in terms in any case?

I thank Professor M. Rees, Dr B. Carr and Professor R. Matzner for constructive comments, and Drs D. W. Sciama and J. J. Binney for reading the manuscript. This work was supported by the SRC.

Received 23 November 1976; accepted 21 March 1977.

- 1 Misner, C. W. *Nature* **214**, 40–41 (1967).
- 2 Partridge, R. B. & Wilkinson, D. T. *Phys. Rev. Lett.* **18**, 557–559 (1967).
- 3 Criss, T. B., Matzner, R. A., Ryan, M. P. & Shepley, L. C. in *GRG 7* (ed. Shaviv, G. & Rosen, J.) (Wiley, New York, 1974).
- 4 Davies, P. C. W. *Nature* **263**, 377–80 (1976).
- 5 Wagoner, R. V. *Astrophys. J.* **179**, 343–360 (1973).
- 6 Rees, M. J. *Commun. Astrophys. Space Phys.* **4**, 179–185 (1972).
- 7 Peebles, P. J. E. & Hauser, M. G. *Astrophys. J. Suppl.* **253** (1974).
- 8 Matzner, R. A. & Misner, C. W. *Astrophys. J.* **171**, 433–445 (1972).
- 9 Rees, M. J. *Phys. Rev. Lett.* **28**, 1669–71 (1972).
- 10 Zeldovich, Y. B. *Mon. Not. R. astr. Soc.* **160**, 1p–3p (1972).
- 11 Liang, E. P. T. *Mon. Not. R. astr. Soc.* **171**, 551–555 (1975).
- 12 Jones, B. T. J. *Rev. mod. Phys.* **48**, 107–149 (1976).
- 13 Harrison, E. R. *Mon. Not. R. astr. Soc.* **148**, 119–30 (1970).
- 14 Barrow, J. D. *Mon. Not. R. astr. Soc.* **178**, 625 (1977).
- 15 Sunyaev, R. A. *IAU Symp.* **63** (ed. Longair, M. S.) (Reidel, Hingham, Massachusetts, 1974).
- 16 Peyrand, J. J. *de Phys. Comptes Rendus*, **39**, 88; 306 (1968).
- 17 Carr, B. J. *Astrophys. J.* **201**, 1–19 (1975).
- 18 Peebles, P. J. E. *Phys. Rev.* **1**, 397–399 (1970).
- 19 Belinskii, V. A., Khalatnikov, I. M. & Lifschitz, E. M. *Adv. Phys.* **19**, 525–573 (1970).
- 20 Feynman, R. F. & Gell-Mann, M. *Phys. Rev.* **109**, 193 (1955).
- 21 Silk, J. *Commun. Astrophys. Space Sci.* **5**, 9–14 (1973).
- 22 Liang, E. P. T. *Phys. Lett.* **51A**, 141–3 (1975).
- 23 Störmer, O. *Gen. Rel. Grav.* **1**, 103–115 (1970).
- 24 Meszaros, P. *Astr. Astrophys.* **38**, 5–13 (1975).
- 25 Misner, C. W. *Astrophys. J.* **151**, 431–457 (1968).
- 26 Zeldovich, Y. B., Sunyaev, R. A. & Zeldovich, Y. B. *I.A.U. Symp.* **63** (ed. Longair, M. S.) (Reidel, Hingham, Massachusetts, 1974).
- 27 Gott, J. R. & Rees, M. J. *Astr. Astrophys.* **45**, 365–376 (1975).
- 28 Dahlbacka, G. N., Chapline, G. F. & Weaver, T. A. *Nature* **250**, 36–37 (1974).
- 29 Doroshkevich, A. G., Zeldovich, Y. B. & Novikov, I. D. *Soviet Phys. JETP* **26**, 408–413 (1967).
- 30 Rees, M. J. & Ostriker, J. P. preprint (1976).
- 31 Rees, M. J., *Mon. Not. R. astr. Soc.* **154**, 187–197 (1971).
- 32 Barrow, J. D. & Matzner, R. A. *Mon. Not. R. astr. Soc.* (submitted).
- 33 Tomita, K. *Prog. theoret. Physics*, **43**, 1503 (1972).
- 34 Matzner, R. A. *Astrophys. Space Sci.* **4**, 459–63 (1969).
- 35 Harrison, E. R. *A. Rev. astr. Astrophys.* **11**, 155 (1973).

A physical interpretation of the Haicheng earthquake prediction

C. H. Scholz*

Earthquake Research Institute, University of Tokyo, Japan

A possible explanation for the successful prediction of the 1975 Haicheng earthquake is that the earthquake was triggered by a deformation front that propagated 1,000 km through NE China at a velocity of about 110 km yr⁻¹. The various phenomena that were used to predict the earthquake can be explained by the deformation front.

ON 4 February 1975, an earthquake of magnitude 7.3 occurred near Haicheng in Liaoning Province, north-east China. The earthquake had been accurately predicted by the Chinese, and thus was the first major earthquake to have been successfully predicted.

The prediction of the Haicheng earthquake took place in four stages—long-range, mid-range, short-range and imminent prediction^{1,2}—each stage representing a narrowing of the prediction window in space and time. Each stage was based on various observations judged to be precursory to an earthquake.

In 1970 a national meeting was held to discuss regions of particular seismic hazard within the People's Republic of China. That meeting followed the occurrence of a series of strong earthquakes that had begun with the two Hsingai earthquakes of 1966 and had culminated with a magnitude 7.4 even in the Pohai Gulf in 1969 (Fig. 1). It was postulated that these large earthquakes were migrating to the north-east, and as a result the area north of the Pohai Gulf, including Liaoning Province, was designated as a region in which a large earthquake, of magnitude 5 to 6, was to be expected in the future. The Jinzhou, Zhuanghe and Yalu River faults, all prominent NNE striking faults in southern Liaoning, were apparently recognised as the major structures along the NE trend of earthquake migration and were singled out for special attention. The result of this long-range prediction was an increase in scientific investigation in Liaoning Province, in particular, releveling along the Pohai coast, installation of new seismic stations and initiation of short level lines across the above mentioned faults.

In June 1974 another meeting was held to review the data collected as a result of the intensified studies made in the preceding 4 yr. At that meeting the prediction time span was shortened, to predict an event of magnitude 5–6 within 1–2 yr in the northern part of the Pohai Gulf. This was based on three main observations: (1) microearthquake activity had increased several-fold, beginning in late 1973, over the entire Liaoning Province. These were not foreshocks, being too widespread. This increase in seismicity was reported for the area enclosed by Shenyang, Tantung and Dairen, a triangular region 200 km on a side. (2) Levelling showed that the greatest gradients in vertical motion for the entire Pohai region occur in the Liaotung Peninsula. (3) Short level lines across the Jinzhou fault at Chin Hsien showed rapid vertical movements beginning in mid to late 1973.

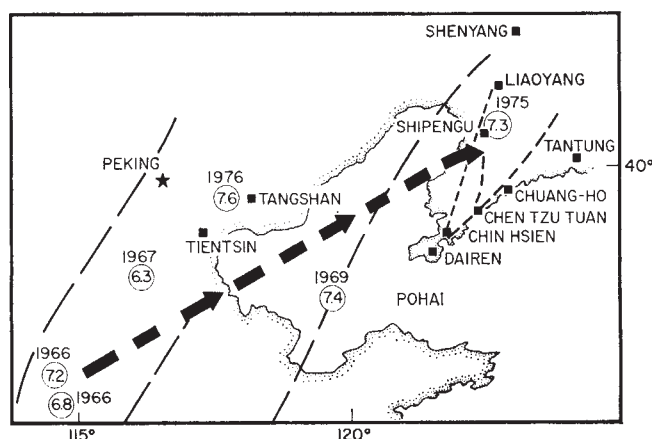
The result of this prediction was to further intensify the scientific efforts, to begin a mass education programme to warn the general populace of the danger of a forthcoming earthquake and to begin organising amateur groups to report anomalous phenomena.

In mid-December 1974, a swarm of earthquakes occurred in the vicinity of Liaoyang, about 70 km north of the epicentre of the Haicheng earthquake. The principal shock was of magnitude 4.8 and was widely felt. This activity is unusual in this region of normally low seismicity, and following these earthquakes the Revolutionary Committee of Liaoning Province issued alerts concerning the possibility of a large future earthquake. Reports of anomalous behaviour of animals and water wells, which had increased before the Liaoyang earthquakes, continued to increase in the region Shenyang–Tantung–Dairen. In mid-January 1975, it was predicted that an earthquake of magnitude 5.5–6 would occur within this region in the first half of 1975. This further shortening of the time span of the prediction (but not of the span in space) seems to have been in response to the Liaoyang earthquakes and the increase in amateur observations of anomalies.

The foreshocks of the Haicheng earthquake began on 1 February, and peaked in the evening of 3 February. Ten of these events were greater than $M=3$, two greater than $M=4$, and were widely felt in the epicentral region. Also beginning on 1 February, observations of anomalous animal behaviour and water-well changes began to migrate to the vicinity of Haicheng and Yinkow (the epicentral region), a tilt anomaly at Shenyang began, and many anomalies in electrical resistivity and radon emanation were observed. These observations, particularly the foreshocks, led to narrowing the prediction area to the Haicheng–Yinkow area on the day of earthquake, and emergency measures were carried out as the earthquake was predicted to occur within 1–2 d.

The earthquake occurred on the evening of 4 February 1975. It was a left-lateral strike-slip event, on a north-west striking fault plane with an aftershock zone about 70 km

Fig. 1 Map of NE China, showing all earthquakes larger than $M=6$ since 1966, the localities mentioned in the text and the path of the deformation front.



*Permanent address: Lamont-Doherty Geological Observatory Columbia University, Palisades, New York 10964, USA.

long³. This fault plane had evidently not been previously recognised, and it has been suggested that it is a newly formed fault⁴.

Discussion

I have attempted briefly to sketch the observations and events that occurred before the earthquake. In one respect or another, most of the observed phenomena can be considered anomalous; what remains to be determined is which of these phenomena were truly precursory to the earthquake. That is, which phenomena had an actual physical connection with the earthquake and could therefore be used reliably to infer the coming of the earthquake.

When I first considered these problems I found that I could, based on conventional seismological thinking, readily dismiss as irrelevant virtually every allegedly precursory phenomenon save the foreshocks. Yet the prediction was made, and accurately, before the foreshocks. That this could have been done based on totally irrelevant observations seems very unlikely, since this would imply that it was fortuitous. But Liaoning Province is an area of normally low seismicity—it seems that in the 3,000 yr of Chinese history, the Haicheng earthquake is by far the largest earthquake ever to have occurred in Liaoning, only two earthquakes greater than magnitude 6 having occurred there in 3,000 yr. It is probably this fact that led the Chinese consistently to underestimate the size of the Haicheng earthquake: at all times they expected it to be less than magnitude 6. Thus barring the possibility of chance, I am forced to consider some rather unconventional possibilities to explain some of the forerunning phenomena.

Among the most interesting of the phenomena is the north-easterly migration of epicentres. That this migration occurred is beyond doubt, but what is important is whether or not there is some causative process responsible for the migration that would allow the Haicheng earthquake to be predicted by induction. Incidentally, the Chinese did not use pure induction. If they had, they would have predicted an earthquake of magnitude 6.5 to 7.5, the size of the other earthquakes in the series, and thereby would have correctly predicted the magnitude of the Haicheng earthquake. Instead, probably for the rather conservative reasons given above, they underestimated its magnitude.

Such migrations are known^{5,6}. One possible explanation for them might be that the change in the strain field produced by one earthquake triggers, with some time delay, the next shock. One cannot explain the north China migration so readily, however. The distance between the successive earthquakes is considerably larger than the fault lengths that can be associated with them, so that the strain change in one epicentral region resulting from the previous earthquakes in the series will be very small. Furthermore, the successive earthquakes did not occur on the same fault.

Two possibilities remain—either the migration was illusory or it occurred as a result of some migrating stress source of presumably deep-seated origin. In what follows, I refer to such a moving stress source as a 'deformation front'.

The second most interesting phenomenon was the strong increase in microearthquake activity that began in late 1973. This increase occurred over the entire province, so it could not have been caused directly by some process in the focal zone of the Haicheng earthquake, nor was it an artefact of increased instrumental coverage.

We are drawn, then, to suspect that if there is some relationship between the increase of microearthquake activity and the ensuing Haicheng earthquake, that it must have been caused by a rapid regional increase of stress level over all of Liaoning Province that later resulted as one of its manifestations in the Haicheng earthquake. This of course is not inconsistent with the earlier suggestion made to explain the earthquake migration, and that late

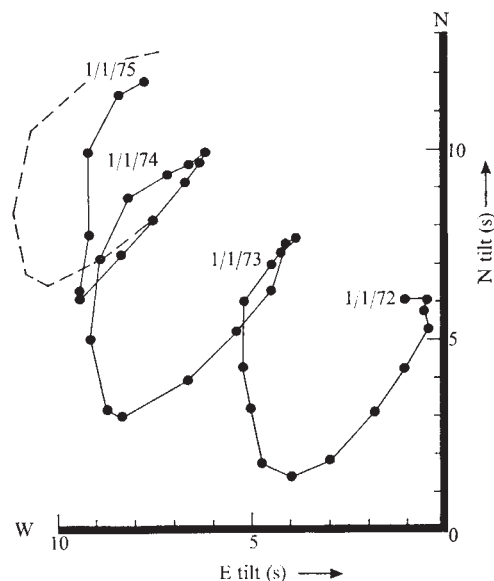


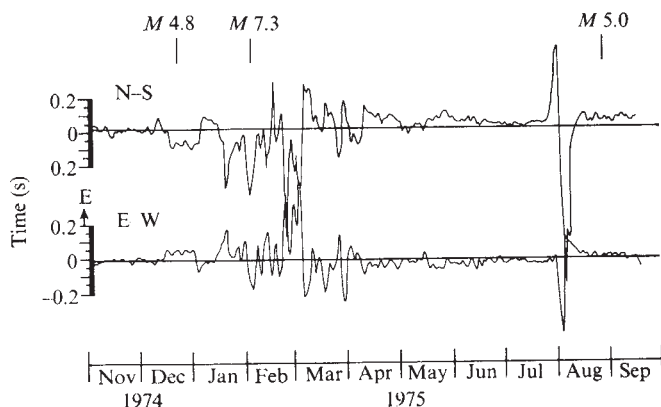
Fig. 2 Monthly tilt recorded at Shipengu. Recorded tilt is given by solid lines and dots. The average annual tilt cycle for 1972 and 1973 is shown as a dashed trace in 1974.

1973 marked the arrival of the stress source in Liaoning Province.

Regional levelling showed that vertical movements had been occurring at a steady rate before 1971 (ref. 2). Short baseline levelling repeated daily at Chin Hsien, however, showed an abrupt onset of tilt, down to the north-west, beginning in mid to late 1973, a time that coincides with the increase in seismicity. This levelling line crosses the Jinzhou fault, but no discontinuity in vertical movements occurred at the fault so that it apparently represents regional deformation. Levelling at Chuang-ho also is consistent with anomalously rapid vertical movements in the period 1970–74, but that at Cheng Tzu Tuan is inconsistent with this. Following the onset of tilt at Chin Hsien, a long-term tilt anomaly was observed with tiltmeters at Shipengu, beginning in June 1974, 8 months before the earthquake (Fig. 2).

Daily tilt rates at Shenyang, farther to the north, are shown in Fig. 3. A short-term anomaly was recorded with these instruments for 4 d before the earthquake. That anomaly was probably not due to deformation within the epicentral region, however, because no such anomaly was

Fig. 3 Daily tilt rates at Shenyang. The baseline assumed for integration is given by the horizontal lines. The integration was only carried to July because of the large tilt anomaly at the beginning of August, which was apparently caused by heavy rainfall.



observed at Shipengu, only 20 km from the epicentre, and no coseismic tilt was recorded at Shenyang. We notice a long-term tilt anomaly at Shenyang, however, marked by the very high tilt rates beginning in mid-December 1974, and lasting until April or May of 1975.

The three tilt anomalies are compared in Fig. 4, where the magnitude of the tilt vector is plotted against time, and the time scales are shifted so that the onset of the anomalies coincide. The time scales in Fig. 4 for Shipengu and Shenyang were expanded by factors of 3.8 and 5.7, respectively, with respect to Chin Hsien, and the magnitude reduced by a factor of four at Shenyang. The anomaly at Shipengu was taken as the difference of tilt recorded in 1974 from the average annual cycle of the previous 2 yr, indicated as the dashed trace in Fig. 2. The Shenyang anomaly was obtained by taking 10-d averages from the data shown in Fig. 3 and integrating them. A steady north-west drift was assumed to be the secular trend there, and was used as the baseline for integration. This procedure will result in the Shenyang tilt anomaly being contaminated with an unknown annual term, but the very good correlation with tilt at Chin Hsien (Fig. 4) shows that the anomaly is not masked. The details of the correlation are lost at Shipengu, where we have only monthly averages, but the overall shape of the anomaly there is also very similar to the others.

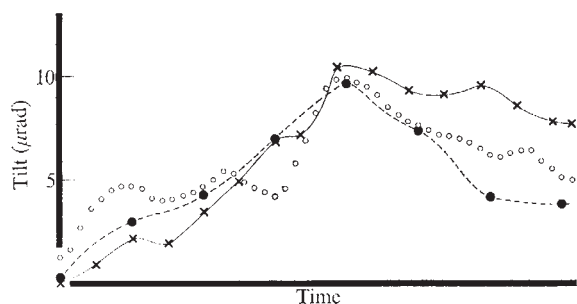


Fig. 4 Comparison of the tilt anomalies at three sites. The time scales at Shipengu (●) and Shenyang (×) were expanded by factors of 3.8 and 5.7, respectively, to that at Chin Hsien (○). The magnitude of the tilt vector is plotted. The magnitude of the anomaly at Shenyang has been reduced by a factor of 4.

The strong correlation of these three anomalies suggests that they are measures of the same phenomenon, a rapid regional deformation that propagated through Liaoning Province in a north-easterly direction.

Returning to the original hypothesis, if we assume that a deformation front propagated through China on a path indicated as the heavy dashed line in Fig. 1, we can show the sequence of events as a space-time diagram projected along that line (Fig. 5). There we can see that a deformation front moving at a velocity of about 110 km yr^{-1} can account for the increase of seismicity in Liaoning Province, the successive onset of tilt anomalies at Chin Hsien, Shipengu and Shenyang, and the triggering of the Liaoyang swarm and the Haicheng earthquake.

The details of the events in Liaoning Province are shown in a similar space-time plot in Fig. 6, which shows the magnitude of the tilt anomalies and, as arrows, the direction of tilt at several times during the passage of the deformation front. Initially, tilt was NNW at Chin Hsien, SSE at Shenyang and east at Shipengu. As the deformation front passed, all tilts rotated clockwise to the west. The Liaoyang swarm occurred just as the maxima arrived at that point, but the Haicheng earthquake occurred after the maxima had passed.

The compressed time scales at Shipengu and Shenyang relative to Chin Hsien indicate that the direction of motion

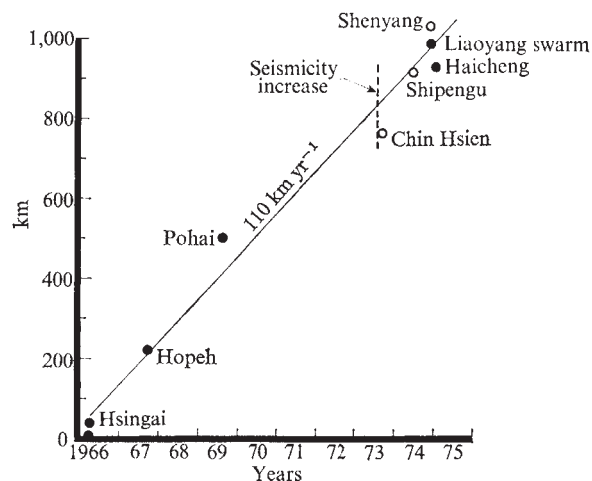


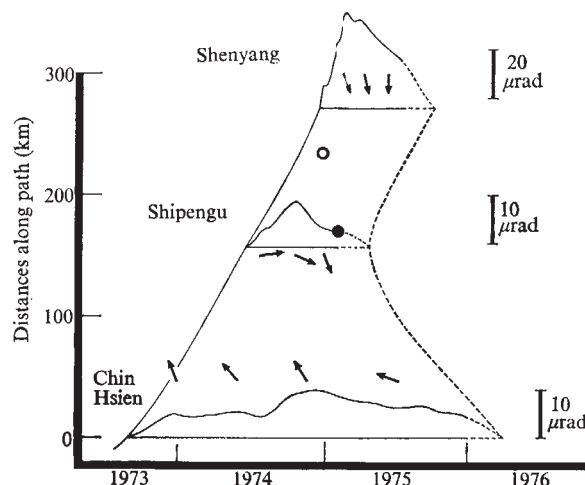
Fig. 5 A space-time diagram projected on a $N60^\circ E$ line.

of the deformation front turned to the east as it entered Liaoning Province in a way suggested in Fig. 1. This rotation was about an axis south of Chin Hsien, so that the propagation velocity of the front was greater by a factor of 3.8 at Shipengu and 5.7 at Shenyang than that at Chin Hsien. This also explains why the anomalies at Shipengu and Shenyang began after, but ended before, the anomaly at Chin Hsien and why the apparent velocity between these sites (200 km yr^{-1}) was greater than velocity deduced from the earthquake migration. Note also that the width of the deformation front was at least 300 km, and that this would include all the earthquake epicentres along the propagation path (Fig. 1).

Two other anomalous phenomena may also have been caused by the deformation front: a rapid rise in sea level of about 1 cm recorded by tide gauges on the Liaoning coast late in 1973, and an increase in magnetic field intensity in Liaoning relative to Peking that began in late 1973 or early 1974 and lasted until late 1975 (refs 1, 2).

There are only two inconsistencies, the lack of increased tilt at Chen Tzu Tuan, in the period 1970–74, and the 1976 Tangshan earthquake, which according to this model should have occurred late in 1968. The first can be explained if the levelling was done early in 1974 at Chen Tzu Tuan.

Fig. 6 Space-time diagram for Liaoning projected on a $N60^\circ E$ line. The magnitude of the tilt anomalies are shown against time, and the direction of the tilt vector indicated by arrows for several periods of time during the anomalies (up is north). The Liaoyang swarm is indicated by an open circle and the Haicheng earthquake by a closed circle.



With regard to the Tangshan earthquake, it should be pointed out that there are two conditions necessary for an earthquake to be triggered by a deformation front. The first is that a zone of high stress, in which an earthquake is already nearly imminent, lies within the path of the deformation front. The second is that the additional stress produced by the deformation front be additive (tensorially) to the pre-existing tectonic stress. The second condition may not have been met at Tangshan.

Conclusions

Although the Chinese did not use the anomalies at Shipengu and Shenyang in their prediction, I have shown that four of the observations upon which they based their predictions—the migration of epicentres, the increase of seismicity in Liaoning, the tilt at Chin Hsien and the Liaoyang swarm—may all be explained by a deformation front that later triggered the Haicheng earthquake.

I have, therefore, proposed a plausible hypothesis for connecting the phenomena used to make the prediction and the occurrence of the earthquake, the success of the prediction being the strongest argument for this plausibility. In doing so I have proposed the existence of a previously unknown phenomenon, the underlying mech-

anism of which is not understood. Deformation fronts have twice been directly observed to propagate between distant strain and tiltmeters in Japan^{7,8}, however. Migrations of large earthquakes, commonly at velocities between 80 and 270 km yr⁻¹ (ref. 6), may be caused by propagating deformation fronts, which may result from mantle stress waves^{9,10}. If this phenomenon is common, it may prove a powerful tool for earthquake prediction.

Much of this work was performed during a visit to China under the auspices of the Committee for Scholarly Communication with the People's Republic of China of the US National Academy of Sciences. The work was supported by NSF, NASA and a Sloan Fellowship. I thank K. Jacob, P. Richards, K. Mogi and K. Kasahara for reviews.

Received 20 December 1976; accepted 10 March 1977.

- ¹ Chu, F. M. *J. Seism. Soc. Japan (Zisin)* 15–26 (1976).
- ² Haicheng Earthquake Study Team, *EOS* (in the press).
- ³ Hsu, S. H. *J. Seism. Soc. Japan (Zisin)* 27–41 (1976).
- ⁴ Wang, T. *et al. Acta geophys. sin.* 3, 204–212 (1976).
- ⁵ Ambraseys, N. N. *Tectonophysics* 9, 143–165 (1970).
- ⁶ Mogi, K. *Bull. Earthquake Res. Inst. Univ. Tokyo* 46, 53–74 (1968).
- ⁷ Kasahara, K. *Phil. Trans. R. Soc. A274*, 287–296 (1973).
- ⁸ Research group for crustal movement, Tohoku Univ. *Symp. Recent Crustal Movements, J. geod. Soc. Japan* (in the press).
- ⁹ Savage, J. C. *J. geophys. Res.* 76, 1954–1966 (1971).
- ¹⁰ Bott, M. H. P. & Dean, D. S. *Nature* 243, 339–341 (1973).

Control of grip and stick in cell adhesion through lateral relationships of membrane glycoproteins

David A. Rees, Clive W. Lloyd & David Thom

Biosciences Division, Unilever Research, Colworth/Welwyn Laboratory, Sharnbrook, Bedford, UK

Although cell adhesion to substrate must ultimately depend on 'sticking' by physical forces at the outer surface, we suggest that control is exercised through the 'grip' of the cytoskeleton. Our experiments indicate that grip can be relaxed by agents which release surface glycoproteins from some form of side-to-side organisation, and can be consolidated by suitable cross-linking to augment this organisation.

ONE of many stimulating challenges to our understanding of the function of animal cells, is to establish the mechanisms by which perturbations at the cell surface are expressed as changes in the cytoskeleton and vice versa. We show here that the response of fibroblast cells to contact with an inert substratum provides a convenient and appropriate system for investigation. When suspended cells settle on a favourable surface, the round shape undergoes a remarkable transformation to a more complex angular form which is flattened against the surface^{1–4}. This involves the formation of a three-dimensional network of fibrillar proteins that is organised in relation to areas of adhesion^{3–5}, so that external contact has triggered a change from an unstructured to a highly organised cytoskeleton. The converse process of detachment from a substrate, for example by trypsin or chelating agents, is equally revealing because the sequence of intermediate stages shows that the primary event is usually the disintegration of a structured cytoskeleton^{6,7}. The cell body rounds up and withdraws from the substrate, leaving sites of adhesion intact; the process does not start

with dissolution of 'the glue' because only later does the cell float away. Further evidence that cell adhesion can be influenced (strengthened or weakened) by perturbing inner as well as outer structures, is that detachment is retarded when the intracellular level of cyclic AMP is artificially raised⁸ to enhance the stability of cytoskeletal elements^{9,9}, and is accelerated by drugs which destabilise these elements^{6,10}. Thus it is important to distinguish between contributions to cell adhesion from internal structures, which we refer to as 'grip', and those from external structures which we refer to as 'stick'. Included in the 'stick' component are all manifestations of physical attraction between the cell surface and the substrate.

Preliminary characterisation of stick and grip

To investigate further whether fibroblast cells are detached from serum-coated glass by mechanisms that involve primarily the lesion of stick, or primarily the lesion of grip, we first used the interference reflection microscope. This instrument^{12–14} can visualise the zones of closest contact between the cell and substrate. These zones at first form a sharply defined pattern of subunits, each about 1–5 μm in area (Fig. 1c, f). Detachment by warm EGTA leaves the zones well defined (Fig. 1f, g) as the cell retracts and the membrane shows pronounced motility and blebbing. We interpret this as a disruption of cytoskeletal stiffening, so that the cell can now peel from the substrate even though the configuration is unaffected at each area of actual contact. From the theory of physical adhesion¹⁵ we know that, other things being equal, a deformable cell would be detached more readily than a stiffened cell because a

deformable body can detach by a peeling mode which is not open to a stiff one. Possible mechanisms that could cause peeling are the cell's own tension on individual sites (shown for example by the ability to deform flexible substrates¹⁶ and by the sharp recoil when adhesions are broken

by micromanipulation⁷), pinching off the adhesive plaques in an exocytotic-like process (we have seen such plaques left behind by cells, remaining attached and enclosed in membrane vesicles) and/or through any turbulence in the medium. Turbulence should also be more effective in detaching deformable cells because they generally present a higher profile to the shearing liquid.

The mechanism of detachment by trypsin is a little more complicated. As the cell slowly rounds under the action of trypsin, it leaves retraction fibrils to the attachment zone^{6,7}. We have found by interference reflection microscopy (Fig. 1c, d, e) that the zones themselves become less well defined and eventually merge into a continuous grey background. Although this might seem to suggest that trypsin has dissolved 'the glue' at each area of contact, such an explanation would be insufficient for the results of further experiments. Using procedures described in the legend to Fig. 2, we found that less than 10% of 16C cells were detached from a monolayer on glass by incubation with trypsin ($10 \mu\text{g ml}^{-1}$) on ice for 40 min. If the trypsin was replaced at 25 min by soybean trypsin inhibitor ($200 \mu\text{g ml}^{-1}$), however, and after a further 10 min in ice, allowed to warm to 37°C , almost all the cells detached immediately. Cells did not detach when cooled in the absence of trypsin and then warmed. Therefore this mild trypsinisation can cause the lesion that normally leads to detachment but this lesion is only expressed as a result of a cellular response made in the absence of trypsin and which requires metabolic activity and/or membrane fluidity. We suggest that trypsin impairs the cell's ability to hold surface macromolecules in an appropriate configuration at each area of contact—hence these configurations disperse and the organised adhesion is replaced by a more amorphous contact over a larger area.

Our provisional conclusion is therefore that adhesion is controlled by cellular activity (that is, by grip) in at least two ways (1) by the overall stiffness and shape which determines how easily the cell is peeled away and (2) by holding macromolecules in the appropriate organised configuration at areas of contact with substrate. We refer to these controls as 'bracing grip' and 'focal grip' respectively. EGTA seems to cause detachment by an action on the former, whereas trypsin has an action on both.

In passing, we point out (as others have^{9,11}) that conditions can be devised for detachment by lesion of 'stick' as well as 'grip'. Cells detached with cold EDTA or cold trypsin in more severe conditions than described above, can have a flat and angular form which is different from the round shape of those detached under conditions more typical of subculturing procedures (Fig. 1a, b). This shows that 'stick' between cell and substrate can sometimes be dissolved or broken by shear before the full alterations in the cytoskeleton which would lead to cell rounding.

Surface glycoproteins in the expression of grip

To investigate further the mechanisms of lesion of fibroblast grip to the substrate, we used plant lectins. These di- or multivalent carbohydrate-binding proteins are known to cross-link carbohydrate components of the outer cell surface, including those glycoproteins which pierce the membrane and are associated with the cytoskeleton. Much evidence^{5,17,18-20} shows that the movement of lectin receptors can be accompanied and sometimes even directed by changes in the cytoskeleton. Our results lead to a conclusion that is opposite to this but complementary—that constraints on lectin receptors can consolidate the cytoskeleton in a form that preserves grip to the substrate.

Several authors^{17,22-25} have previously stated or implied that lectins can inhibit the detachment of cells by proteases and chelating agents. Lectins which apparently have this property are concanavalin A (con A)^{22,23,25}, phytohaemagglutinin^{17,22}, and the lectin of higher molecular weight

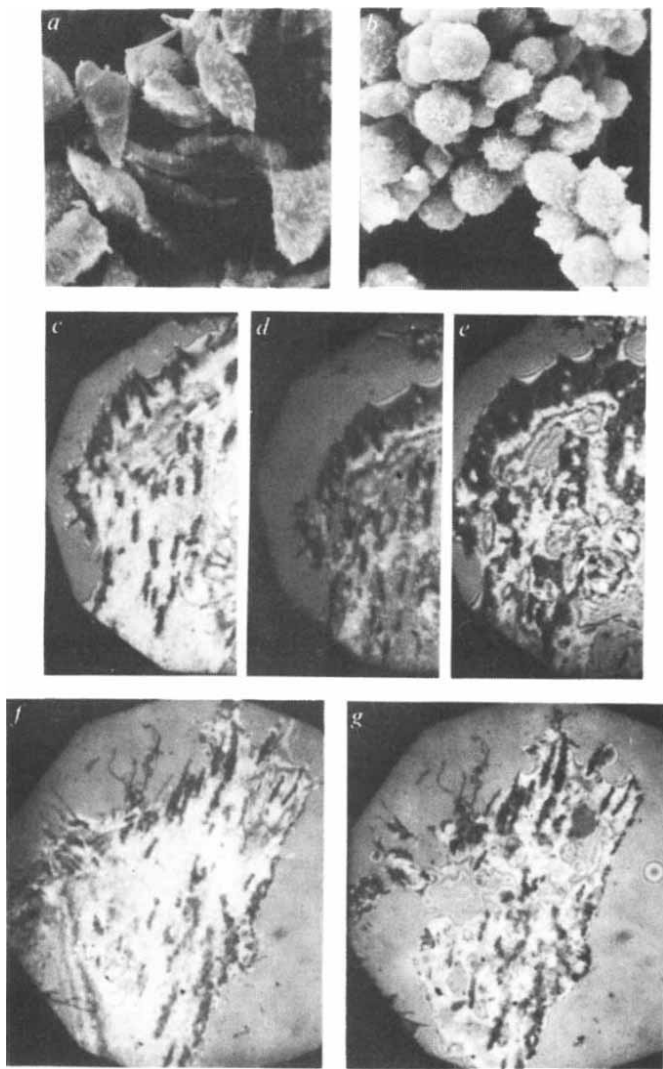


Fig. 1 Photomicrographs of rat dermal fibroblasts (16C line, Colworth strain) as they are influenced by various conditions of detachment from a glass substrate. *a*, Scanning electron micrographs ($\times 952$) of cells detached from a glass medicine bottle after decanting the medium, washing twice with ice-cold Hank's balanced salts solution (HBSS, Ca^{2+} and Mg^{2+} free), chilling in this solution for 15 min, then shaking off with ice-cold EDTA (0.02% in the same salt solution). Cells were immediately fixed by mixing with an equal volume of 3% glutaraldehyde in Dulbecco's phosphate-buffered saline (PBS). After 30 min at 0°C , the cells were centrifuged, washed twice with PBS, post-fixed with osmium tetroxide and prepared for microscopy by critical point drying. *b*, Cells as they appear after typical dissociation with trypsin (0.05%)–EDTA (0.01%) in PBS, or with EGTA (0.05%) in Tris-buffered Ca^{2+} and Mg^{2+} free saline, at room temperature. Washing, fixing and processing as above. *c, d, e*, Stages in detachment by trypsin ($100 \mu\text{g ml}^{-1}$ in HBSS at 23°C) from a glass coverslip, as seen with interference reflection optics. Parts of the cell that are in close apposition to the glass surface (that is, the adhesion sites) appear as dense black zones. *c*, Before exposure to trypsin; *d*, after exposure for 7 min; *e*, after exposure for 15 min. Note that attachment sites are at first sharply defined and later spread and become continuous. For space reasons, only a part of the underside of one cell is shown. *f, g*, Stages in detachment by ethylene glycol bis (β -aminoethyl ether)-*N, N*-tetraacetic acid (EGTA, 0.05% in Ca^{2+} and Mg^{2+} free HBSS at 37°C). *f*, Before exposure to EGTA; *g*, after exposure for 35 min. Note that the size, shape, and mutual relationship of many attachment sites alter little while an area of the cell on the left of the photograph peels away from the substrate.

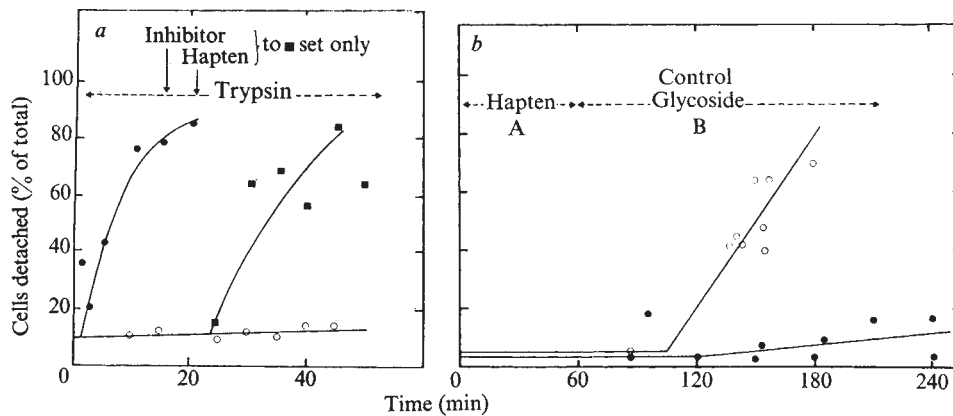


Fig. 2 Rate of detachment of fibroblasts from a glass substrate by trypsin, and its modification by con A. Established monolayers in glass Leighton tubes (without coverslips) were gently treated with the sequence of solutions indicated below. Each point of the graph represents the proportion of cells which could be detached (Coulter assay) by gently rolling the medium several times over the cells in a particular tube at a particular time, essentially as described by Johnson and Pastan⁸. *a*, Results for three sets of tubes of NIL 8 cells. Each tube of the first set (●), was gently washed with HBSS, incubated at 37 °C for 20 min with the same solution, then at 37 °C with crystalline trypsin (25 $\mu\text{g ml}^{-1}$) in HBSS. Each tube of the second set (○), was gently washed with HBSS, incubated at 37 °C for 20 min with con A (100 $\mu\text{g ml}^{-1}$) in HBSS, then with trypsin as for the first set. Each of the third (■), was treated in the same way as the tubes of second set except that, after 15 min in trypsin, the cells were washed and incubated first with trypsin inhibitor (from soy bean, 100 $\mu\text{g ml}^{-1}$ in HBSS; 5 min at 37 °C) then with methyl α -D-mannoside (0.05 M in HBSS at 37 °C). All these operations and the results are shown schematically in Fig. 3. *b*, 16C cells were washed with HBSS, incubated at 37 °C with con A (100 $\mu\text{g ml}^{-1}$ in HBSS at 37 °C for 20 min), washed with HBSS and incubated with trypsin (25 $\mu\text{g ml}^{-1}$ in HBSS at 37 °C for 25 min), washed and flooded with trypsin inhibitor (from soy beans, 100 $\mu\text{g ml}^{-1}$ in HBSS) and then placed on ice for 10 min. The tubes were divided into two sets in this stage. One of the first set (○) were washed and incubated with ice-cold methyl α -D-mannopyranoside (0.05 M in HBSS). Tubes of the second (●) were washed and incubated with methyl β -D-galactoside (0.05 M) in HBSS. After 60 min, all cells (both sets) were carefully washed and flooded with ice-cold galactoside solution then allowed to warm to 37 °C. Note that the cold incubation with mannoside is necessary for subsequent detachment of cells in the warm with galactoside. In other control experiments, briefer treatment with mannoside (up to 30 min in the cold) was found to be insufficient for rapid detachment of a high proportion of cells when subsequently warmed with galactoside. The significant operations and results are shown schematically in Fig. 3.

from *Ricinus communis*²⁴. This has usually been explained in trivial terms, such as blocking of access to the detaching agent or the formation of lectin bridges between cell and substratum to replace the natural adhesions. We now describe experiments which show that these explanations are unlikely and then explore the implications of our own hypothesis, which has not previously been considered, namely, that adhesion is preserved by influencing the cytoskeleton through its relation with surface receptors that can be cross-linked by lectin.

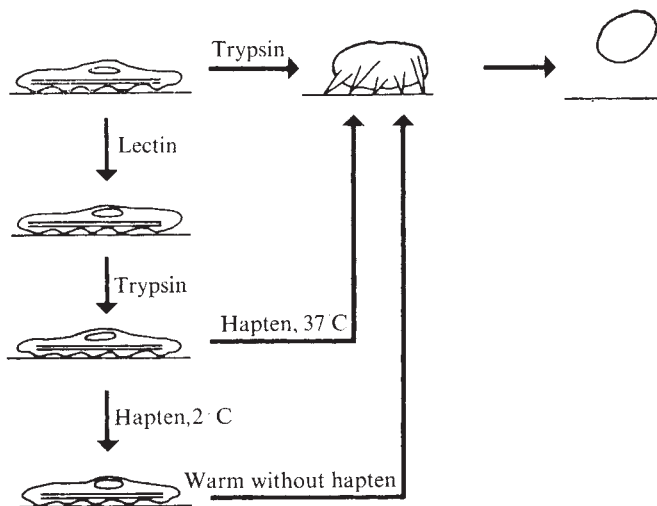


Fig. 3 Schematic representation of key experiments with lectins and cell detachment. The top line shows that cells normally round up with warm trypsin, leaving retraction fibrils which are then withdrawn as the cell separates from the substrate. The vertical sequence shows that pretreatment with con A prevents subsequent rounding and detachment under the influence of trypsin; after removal of the trypsin, however, displacement of con A at 37 °C (see first loop) causes rounding and detachment proving that the lectin block was in the expression of the tryptic lesion rather than at the level of the action of trypsin. The last step in the vertical sequence shows that the displacement of lectin from trypsinised, protected cells is not in itself sufficient to cause detachment—subsequent warming is necessary to re-start metabolic activity and loosen the grip.

Exposure of cells in monolayer to trypsin caused rapid rounding and detachment but when they had been pre-incubated with con A, very few were detached by this treatment. Even vigorous shaking and further amounts of enzyme produced little further effect. Measurements with ¹²⁵I-labelled con A showed that 16C cells bound 7.2×10^7 molecules of lectin per cell, most of which (5.7×10^7 molecules per cell) was inhibited in the presence of 0.05 M α -D-mannopyranoside. Trypsinisation (25 $\mu\text{g ml}^{-1}$ for 20 min at 37 °C) removed 32% of the specifically bound lectin, presumably as a combined result of cleavage of the lectin itself and of a proportion of the receptors for it. The con A-protected trypsinised cells could be detached readily by brief incubation with methyl α -D-mannoside (Fig. 2a). Since this effect was specific in that it was not observed with methyl β -D-galactoside, it is attributable to the displacement of the remaining bound con A and hence of its protective influence. These observations (shown schematically in Fig. 3) prove that target molecules have been split by trypsin in the usual way but that detachment was prevented by the bound lectin; the explanation of lectin protection in terms of mere blockage of the action of trypsin, is therefore untenable. Similar results but with interesting differences in detail (unpublished work in this laboratory), were obtained with EGTA detachment. Evidently both agents can act on lectin-coated cells to cause whatever lesion would normally result in cell detachment but the usual expression of each lesion is prevented until con A is displaced. We have observed that some (but not all) other lectins can act analogously to con A. The effect is observed with all the normal and transformed cell lines that we have studied.

After the action of trypsin, cell attachment is not retained merely through lectin bridges to the substrate (Fig. 3). If this were so, removal of these bridges would leave cells free to be washed away—whereas we find that they are still firmly attached after prolonged incubation with methyl α -D-mannoside in the cold (Fig. 2b). Measurements with ¹²⁵I-con A confirmed that the cold incubation was effective in displacing 62% of the bound trypsin-resistant lectin, which is evidently sufficient to allow the cells to detach when metabolic activity is allowed to re-start by subsequent warming to 37 °C in the absence of hapten (Fig. 2b). Cells did not detach extensively after shorter incubations with cold mannoside followed by warming after its removal (for

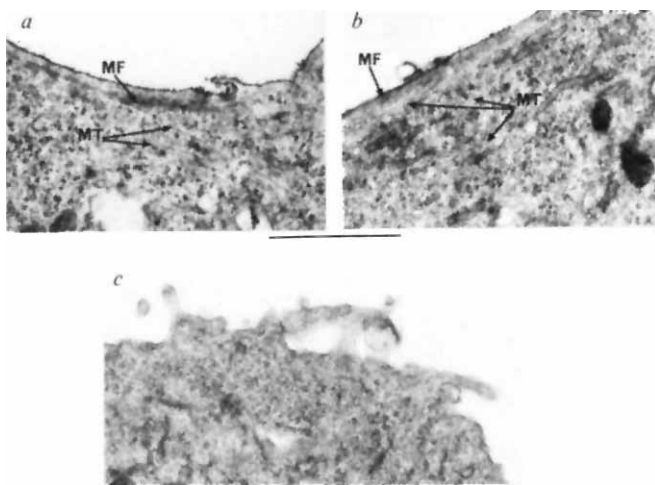


Fig. 4 Typical electron micrographs of BHK cells seeded singly and grown on melanex coverslips for 48 h. After appropriate treatment (see below), cells were fixed for 30 min at 37 °C in 3% glutaraldehyde with 4% tannic acid in 0.1 M cacodylate buffer pH 7.3, then washed briefly in 0.1 M cacodylate, and postfixed with 1% osmium tetroxide in 0.1 M cacodylate pH 7.3 for 30 min at room temperature then 1% aqueous uranyl acetate for 30 min. Dehydration was through a graded series of ethanol concentrations. The sample was embedded in Araldite and sections were cut parallel to the substrate and stained with lead citrate. *a*, Control BHK cell. *b*, BHK cell after incubation at 37 °C with con A (100 $\mu\text{g ml}^{-1}$ in HBSS) for 20 min followed by trypsin (25 $\mu\text{g ml}^{-1}$ in HBSS) for 30 min. *c*, BHK cell after incubation at 37 °C with trypsin (25 $\mu\text{g ml}^{-1}$ in HBSS) for 10 min. MT, microtubules; MF, microfilament bundles. Bar 1 μm .

example, after 30 min which removed only 30% of bound trypsin-resistant lectin).

Having discounted two explanations, we are left with the possibility that lectin constrains the lateral arrangement of glycoproteins in a way that preserves the integrity of cytoskeletal components responsible for grip. The rapid reversal of protective action when lectin is displaced in our

experiments by glycoside hapten confirms that effective binding is indeed on the outer membrane surface, rather than some result that follows endocytosis.

The cytoskeleton in the expression of grip

To characterise the response of the cytoskeleton to detachment agents and lectins, we followed changes in BHK cells that had been seeded singly and allowed to spread for 48 h. Thin sections cut parallel to the substrate showed that areas of spread membrane contained long microtubules and bundles of microfilaments close to the underside of the cell (Fig. 4), and were bounded by a stretched, linear cell margin. By contrast, light microscopy showed trypsinised cells to be rounded and blebby, and thin sections confirmed this loss of surface tone. The convolution and infolding of the cell's margin is accompanied by progressive fragmentation and the eventual disappearance of microfilament bundles and microtubules prior to cell detachment (Fig. 4). When cells had been pretreated with con A, however, these trypsin-induced changes were arrested at an early stage (Fig. 4). Although cell shape may be altered, the margins remained generally smooth, long microtubules remained in most cells, and microfilament bundles could be observed in many.

From these experiments we can therefore confirm that pretreatment with con A prevents the action of trypsin from causing not only cell detachment but also the dissolution of cytoskeleton. There is some indication that con A might be more efficient in arresting the dispersal of microtubules than of microfilament bundles but the evidence for this is not clear cut. For the time being at least we prefer to regard these components as linked in a single system which is perturbed together by the agents we have used.

Model for glycoprotein relationships in stick and grip

We conclude that lectin cross-linking of carbohydrate chains on the outer surface can constrain the cell in a way that keeps it flattened and attached to the substrate, despite the action of agents that normally cause it to withdraw and detach. Whether the target sites for these agents are inside or outside the cell, the lesions they induce are not expressed until after the receptors have been released to rearrange in the plane of the membrane. At least two types of lesion can be involved in cell detachment: dispersal of the adhesive areas, for example by trypsin, and loosening of the cell body without changing the form and distribution of these areas, for example by EGTA. In our terms these represent lesions in focal grip and bracing grip respectively. The expression of both can be blocked by appropriate lectins.

We now propose a model to account for these results. In the suspended cell, surface receptors are not organised (Fig. 5, non-grip configuration). The development of the stiff, flat and attached form of the cell, however, requires lateral organisation of surface glycoproteins for the functions of focal grip and bracing grip (Fig. 5). Note that this representation is highly schematic and does not attempt to show the relative dimensions of molecular components of the surface and cytoskeleton (compare with ref. 26) or indeed to distinguish between different components of the cytoskeleton such as microfilament bundles and microtubules.

The relevant glycoproteins are exposed on the outer surface and pass through the lipid membrane to integrate with the cytoskeleton in such a way that ordering of the cytoskeleton necessarily involves an ordering of the lateral arrangement of these glycoproteins and likewise for dis-ordering. In the morphologically developed cell, specific side-to-side organisation of these glycoproteins is presumably maintained by natural bridging molecules, inside and/or

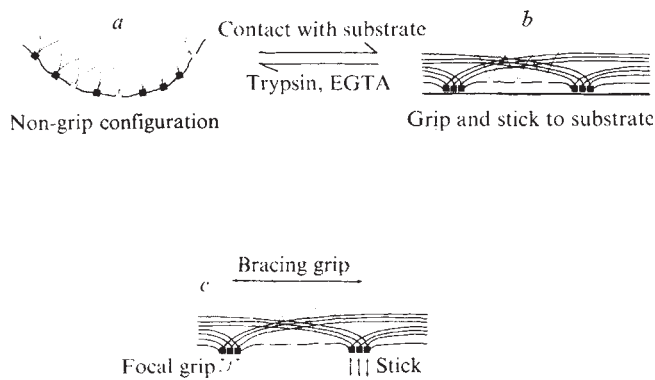


Fig. 5 Highly schematic model for lateral organisation of cell surface glycoproteins for grip and stick in cell adhesion to substrate—note the reservations in the text on molecular realism. *a*, The suspended cell is not organised intra- or extracellularly for adhesion. *b*, *c*, In the attached cell, surface molecules (■) have been moved and organised at the sites of adhesion in a way that sets up the development of the cytoskeleton; they are held in this configuration by focal grip. The full development of cell morphology requires further surface molecules (□) to be organised at sites outside the adhesions, in the process by which cell membrane is drawn down between adhesions to flatten the cell and stiffen it against peeling, that is, to maintain bracing grip. Cell rounding and detachment can be initiated by lesions in the natural side-to-side associations of either set of surface glycoproteins (■ or □); expression of either lesion is prevented by lectins which, by cross-linking the glycoproteins, suitably constrain their positions.

outside the membrane—and these are targets for agents which cause detachment. Such bridging molecules have yet to be characterised but submembranous, fibrillar proteins and external, endogenous, trypsin-sensitive, agglutinins^{28,29} could clearly be involved.

We thank Mrs T. Ogden for skilled technical support and Messrs C. G. Smith, M. Stubbs, C. Allcock and F. J. Judge, and Dr G. A. Dunn for help and advice with the microscopes.

Received 30 December 1976; accepted 20 March 1977.

- ¹ Witkowski, J. A. & Brighton, W. D. *Expt Cell Res.* **68**, 372–380 (1971); **70**, 41–48 (1972).
- ² Rajaraman, R., Rounds, D. E., Yen, S. B. S. & Rembaum, A. *Expt Cell Res.* **88**, 327–339 (1974).
- ³ Abercrombie, M., Heaysman, J. E. M. & Pegrum, S. M., *Expt Cell Res.* **67**, 359–367 (1971).
- ⁴ Pegrum, S. M. & Maroudas, N. G. *Expt Cell Res.* **96**, 416–422 (1975).
- ⁵ Goldman, R. D. in *Locomotion of Tissue Cells* (Ciba Foundation Symp. 14, new series), 83–107 (Elsevier, Amsterdam, 1973).

- ⁶ Revel, J. P., Hoch, P. & Ho, D. *Expt Cell Res.* **84**, 207–218 (1974).
- ⁷ Harris, A. *Devl Biol.* **35**, 97–114 (1973).
- ⁸ Johnson, G. S. & Pastan, I. *Nature new Biol.* **236**, 247–249 (1972).
- ⁹ Willingham, M. C. & Pastan, I. *J. Cell Biol.* **67**, 146–159 (1975).
- ¹⁰ Shields, R. & Pollock, J. *Cell* **3**, 31–38 (1974).
- ¹¹ Vasiliev, Y. M., Gelfand, I. M. & Tint, L. S. *Tsitologiya* **17**, 633–638 (1975).
- ¹² Curtis, A. S. G. *J. Cell Biol.* **20**, 199–215 (1964).
- ¹³ Abercrombie, M. & Dunn, G. A. *Expt Cell Res.* **92**, 57–62 (1975).
- ¹⁴ Izzard, G. S. & Lochner, L. R. *J. Cell Sci.* (in the press).
- ¹⁵ Good, R. J. in *Treatise on Adhesion* (ed. by Patrick, R. L.) 9–68 (Dekker, New York, 1967).
- ¹⁶ Harris, A. K. in *Locomotion of Tissue Cells* (Ciba Foundation Symp. 14, new series), 3–20 (Elsevier, Amsterdam 1973).
- ¹⁷ Gail, M. in *Locomotion of Tissue Cells* (Ciba Foundation Symp. 14, new series), 287–310 (Elsevier, Amsterdam, 1973).
- ¹⁸ Ukena, T. E., Borysenko, J. Z., Karnovsky, M. J. & Berlin, R. D. *J. Cell Biol.* **61**, 70–82 (1974).
- ¹⁹ Poste, G., Papahadjopoulos, D. & Nicolson, G. L. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4430–4434 (1975).
- ²⁰ Berlin, R. D., Oliver, J. M., Ukena, T. E. & Yin, H. H. *Nature* **247**, 45–46 (1974).
- ²¹ Ji, T. H. & Nicolson, G. L., *Proc. natn. Acad. Sci. U.S.A.* **71**, 2212–2216 (1974).
- ²² Grinnell, F. *J. Cell Biol.* **58**, 602–607 (1973).
- ²³ Sato, C. & Takasawa-Nishizawa, K. *Expt Cell Res.* **89**, 121–126 (1974).
- ²⁴ Gahmberg, C. G. & Hakomori, S. *J. biol. Chem.* **250**, 2447–2451 (1975).
- ²⁵ Shoham, J., Inbar, M. & Sachs, L. *Nature* **227**, 1244–1246 (1970).
- ²⁶ Loor, F. *Nature* **264**, 272–273 (1976).
- ²⁷ Hynes, R. O. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3170–3174 (1973).
- ²⁸ Yamada, K. M., Yamada, S. S. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3158–3162 (1975).

letters to nature

Cosmological tests of general relativity

GENERAL relativity theory could be tested on a cosmological scale by measuring the Hubble constant and the deceleration parameter, if, in addition, everything could be known about the matter filling the universe. If, on the other hand, nothing could be presupposed about the matter content of the universe, general relativity could not be tested by measuring any number of time derivatives of the scale factor. But upon making the assumption of a universe filled with a non-interacting mixture of non-relativistic matter and radiation we can in principle test general relativity by measuring the first five derivatives of the scale factor. In the following, some general relations are presented using this assumption.

A reasonable model for the universe in a not too early stage of its development is the Robertson–Walker metric

$$ds^2 = -dt^2 + R^2(t) [(dr^2)/(1 - kr^2) + r^2 d\Omega^2] \quad (1)$$

describing a homogeneous isotropic universe.

From general relativity theory it follows that the scale factor $R(t)$ obeys the Friedmann equation

$$R^2 + k = \frac{8\pi G}{3} R \rho + \frac{\lambda}{3} R^2 \quad (2)$$

and conservation of energy and momentum is expressed by

$$\frac{d}{dt} (\rho R^3) = -3pR^2 \quad (3)$$

Here and throughout this paper we put $c = 1$; λ stands for the cosmological constant, while ρ and p denote the density and pressure of the matter content of the universe.

To test these equations we need observations. By measuring the luminosity distance or the angular diameter distance of remote objects as a function of redshift z , we can in principle determine the Hubble constant $H = \dot{R}/R$, the deceleration parameter $q = -\ddot{R}R/\dot{R}^2$, and an indefinite number of similar dimensionless parameters constructed from higher time derivatives¹. We shall use here, following Harrison²

$$Q = \frac{\ddot{R}R^2}{\dot{R}^3} \quad X = \frac{R^{(4)}R^3}{\dot{R}^4} \quad Y = \frac{R^{(5)}R^4}{\dot{R}^5} \quad (4)$$

(all evaluated at the present time, with signs which make them positive for vanishing λ , as will be shown). Without using Friedmann's equation (2), however, knowledge of any number of those parameters will not suffice to obtain the value of the present curvature $K = k/R^2$ (ref. 3). But even equations (2) and (3) are not sufficient without additional information about the equation of state, as can be seen as follows: by differentiating equation (2) N times we obtain $(N + 1)$ equations for $(N + 2)$ unknowns: ρ and its first N derivatives and λ .

A plausible form for the equation of state is obtained by assuming that at present the universe contains a non-interacting mixture of two forms of matter, non-relativistic matter with density ρ_m and ultra-relativistic matter with density ρ_r . Then

$$\left. \begin{aligned} \rho &= \rho_m + \rho_r \\ \rho &= \frac{1}{3} \rho_r \end{aligned} \right\} \quad (5)$$

For the non-relativistic matter, density, particle number conservation gives

$$\rho_m = A/R^3 \quad (6)$$

and thus with (3)

$$\rho_r = B/R^4 \quad (7)$$

where A and B are constant.

Now, carrying out three successive time differentiations of equation (2) leads with (2) to four equations for the newly introduced time derivative parameters. Solving these for the quantities K , λ , ρ_m and ρ_r , gives

$$K = \frac{1}{2} (-X + 4Q + Qq + 2q - 2) H^2 \quad (8)$$

$$\lambda = \frac{1}{4} (-X + 6Q + Qq - 6q) H^2 \quad (9)$$

$$\rho_m = \frac{1}{4\pi G} (-X + 3Q + Qq + 3q) H^2 \quad (10)$$

$$\rho_r = \frac{3}{32\pi G} (X - 2Q - Qq - 2q) H^2 \quad (11)$$

Differentiating (2) a fourth time finally leads to a relation

between the four parameters q , Q , X and Y which should hold if general relativity is valid

$$Y = 7X + 3Xq - 6Q + Q^2 - 7Qq - 3Qq^2 - 6q \quad (12)$$

This condition provides (in principle) a full test of the theory.

If less than five parameters measuring the time derivatives of the scale factor are known—as certainly will be the case in practice—still conditional relations can be derived between the various parameters. From equations (10) and (11) it follows that, since $\rho_m > 0$ and $\rho_r > 0$

$$Qq + 2(Q + q) < X < Qq + 3(Q + q) \quad (13)$$

a condition valid if only q , Q and X are known.

If X is not available either we introduce a new parameter α , still to be able to express the four quantities K , λ , ρ_m and ρ_r explicitly. For α we take

$$\alpha = \rho_r / (\rho_m + \rho_r) \quad 0 \leq \alpha \leq 1 \quad (14)$$

and we then obtain

$$K = [-1 + (3 + \alpha)/(3 + 5\alpha)Q - (4\alpha)/(3 + 5\alpha)q]H^2 \quad (15)$$

$$\rho = 1/8\pi G 6/(3 + 5\alpha)(Q + q)H^2 \quad (16)$$

$$\lambda = 3[(1 + \alpha)/(3 + 5\alpha)Q - (2 + 4\alpha)/(3 + 5\alpha)q]H^2 \quad (17)$$

Equation (16) gives us, as $\rho > 0$, the condition

$$Q + q > 0 \quad (18)$$

If we only know H and q —a situation which is at present striven for—it is regrettably no longer possible to write down even conditional equations, unless observational evidence of a different nature, for example, regarding the age of particular objects in the universe, is taken into account.

Figure 1 illustrates a classification, based on equations (15) and (16), of the structural and evolutionary properties of the universe, as depending on the values of q and Q , but with arbitrary λ . We can distinguish between domains of positive and negative ρ (the latter physically unrealistic), of $k = \pm 1$ (separated by a line $k = 0$), and the presence or absence of an initial and a final singularity (only with a final singularity the universe, at present expanding, has a finite lifetime). Depending on the value of α , the relative amount of matter which is present in the form of radiation, the borderlines between regions may change. Their equations are given in the two extreme cases $\alpha = 0$ and $\alpha = 1$ as

$$Q = q^2 + q + 1; \quad \alpha = 1 \quad (19)$$

$$Q = \frac{1}{3} \left[2q^2 + q + 2 + 2 \left(q^4 + \frac{8}{3}q^3 + 2q^2 - \frac{1}{3} \right)^{1/2} \right]; \quad \alpha = 0$$

Where the status of the universe depends on α , this is indicated by a double label as explained in the caption. Comparable diagrams have been published earlier by Stabell and Refsdal⁴. (As a referee commented, Fig. 3 of ref. 4, which is a σ - q diagram, is topologically very similar to Fig. 1 of this paper, as $\sigma = 4\pi G\rho/3H^2 \sim Q + q$.)

We can further simplify equations (8)–(11) by introducing two transformed time-derivative parameters

$$\left. \begin{aligned} \tilde{Q} &= Q + q \\ \tilde{X} &= X - Qq \end{aligned} \right\} \quad (20)$$

Then the right-hand side of equations (8)–(11) can be seen to

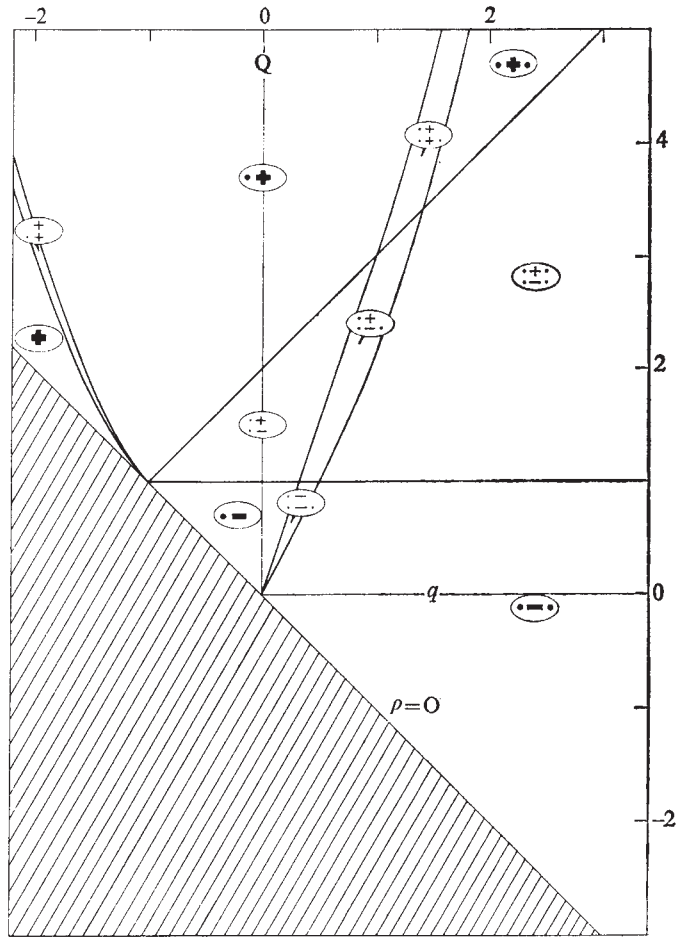


Fig. 1 Classification of structural and evolutionary properties of the universe in terms of Q and q with arbitrary λ . The shaded area is unphysical since $\rho < 0$ there. The symbols in the other regions have the following meaning: if one plus or minus sign is given, the value of the topological parameter k is $+1$ or -1 , respectively. A dot to the left of the sign denotes the presence of an initial singularity, and a dot to the right a final singularity. If a region has only one label, the conclusions follow irrespective of the value of α . If it bears two labels, the upper label denotes the situation for pure non-relativistic matter ($\alpha = 0$) and the lower one applies to the case of pure radiation ($\alpha = 1$). If $0 < \alpha < 1$, as is the case in our universe, the border lines lie always somewhere in between those given by the two extremes $\alpha = 0$ and $\alpha = 1$.

become linear expressions in the new parameters, where equation (12) transforms into

$$Y = 7\tilde{X} + 3\tilde{X}q - 6\tilde{Q} + (\tilde{Q} - q)^2 \quad (12')$$

The conditional relations (13) and (18) become very simple

$$2\tilde{Q} < \tilde{X} < 3\tilde{Q} \quad (13')$$

$$\tilde{Q} > 0 \quad (18')$$

We conclude that surprisingly simple relations can be derived, on the basis of rather general assumptions regarding the structure and matter content of the universe, between the 'physical' parameters of the universe ρ_m , ρ_r and λ , together with K , and the 'geometrical' parameters involving the first four time derivatives of the scale factor. Between the first five time derivatives a strict relation should hold if general relativity applies. This relation forms, in principle, a definite

test for the theory. If less than five parameters are known, but more than two, conditional relations can be derived.

I thank Professor H. G. van Bueren for helpful discussions.

P. HUT

*The Astronomical Institute,
Utrecht,
The Netherlands*

Received 1 December 1976; accepted 22 February 1977.

¹ Weinberg, S. *Gravitation and Cosmology* (Wiley, New York, 1972).

² Harrison, E. R. *Nature* 260, 591 (1976).

³ Weinberg, S. *Astrophys. J.* 161, L233 (1970).

⁴ Stabell, R. & Refsdal, S. *Mon. Not. R. astr. Soc.* 132, 379 (1966).

Tugboat model for OB binaries, X-ray stars and pulsars

AN examination of the kinematical properties of binary OB stars, binary X-ray sources and pulsars suggests an evolutionary sequence linking an apparent low-velocity class of pulsars to the binary nature of their extreme Population I progenitors.

In a compilation of OB supergiants by Hutchings¹, 10 of the 65 stars listed are known binaries, and several more have variable velocities indicating a possible binary nature. Table 1 summarises the kinematical properties of these systems, grouping them according to the nature of their companions—four which have normal stars as secondaries (that is unevolved systems), three which have invisible, low-mass companions and three which are known X-ray sources presumably containing either a neutron star or a black hole. Comparison of columns 3 to 5 of Table 1 reveals distinct kinematical differences among these three groups. The unevolved systems have low space velocities and z -heights typical of extreme Population I objects. The group of stars with unseen, low-mass companions have similar z -heights, but much higher velocities. Their low- z values indicate that these velocities must have been acquired recently, for such speeds would indicate origins far (> 500 pc) from the plane if they were to be maintained on evolutionary timescales ($\sim 10^7$ yr). The velocities for the X-ray sources are not well determined, but their distances from the plane are larger than that for typical extreme Population I objects.

A careful comparison of the z -distribution of pulsars with their characteristic ages given by $\tau = P/2\dot{P}$ yields several clues about the kinematics and evolution of the pulsar population. The pulsar scale height $z_{r.m.s.} \simeq 250$ pc (ref. 8) indicates that these objects must in general have velocities much larger than those of the massive OB stars ($z_{r.m.s.} \simeq 80$ pc) from which they

are thought to originate. Recent measurements of pulsar proper motions^{9–12} have confirmed this conclusion, finding transverse velocities ranging from 50 km s^{-1} to well over 500 km s^{-1} . For $\tau < \sim 5 \times 10^6$ yr, the observed z -distribution, properly corrected for selection effects¹³, shows a rise of scale height with age that is consistent with the observed pulsar velocities. There is, however, a significant fraction of pulsars ($\sim 25\%$) with much larger characteristic ages, whose z -distribution implies absurdly low velocities (often with $v_z \ll 1 \text{ km s}^{-1}$). Proper motion measurements for several of these sources have yielded much higher velocities, showing that the characteristic age is not a true indicator of chronological age after a few Myr, thus confirming such suggestions made earlier from other evidence^{14,15}.

The discrepancy between dynamical and characteristic ages may be related to the decrease in the quantity $P\dot{P}$ ($\propto B^2 R^6 / I$) with time, perhaps as a result of magnetic field decay^{14,15}, alignment of the pulsar's magnetic and spin axes¹⁶, or some other process. One can model such a decay as an exponential decrease in $P\dot{P}$ with a characteristic timescale τ_D given by the time at which the characteristic and dynamical ages diverge ($\tau_D \sim 5 \times 10^6$ yr), and thus derive a new age estimate for each source, as well as the quantity $P_0\dot{P}_0$ proportional to the sources' initial magnetic field strength and moment of inertia. Comparison of $P_0\dot{P}_0$ with measured pulsar velocities shows that the three sources (Class A) with the lowest initial fields have $\langle v_t \rangle_A \lesssim 60 \text{ km s}^{-1}$ while the remaining nine sources (Class B) all have $v_t > 100 \text{ km s}^{-1}$ with $\langle v_t \rangle_B \sim 265 \text{ km s}^{-1}$. Furthermore, from the much larger sample of sources with measured values of $\dot{P}$, we find that the 18 sources with small $P_0\dot{P}_0$ have a mean $|z|$ of ~ 75 pc, while the 59 sources with larger $P_0\dot{P}_0$ have $\langle |z| \rangle_B \sim 250$ pc (ref. 13).

We have proposed an explanation of these two dynamical classes of pulsars in terms of the different environments in which they originate¹³. The high-velocity, widely distributed sources are supposed to originate primarily from single stars. They were presumably accelerated by means of the asymmetric radiation force¹⁷, which provides a natural explanation for their wide range of velocities through small variations in the asymmetry coefficient. The second class of pulsars is presumed to arise in tightly bound, massive binaries where the asymmetric force is insufficient to unbind the system (the supernova explosion that produces the neutron star will not usually unbind the system either¹⁸). The force thus acts on the system as a whole, accelerating it to a velocity that is reduced from that which the free neutron star would have attained by the ratio of the stellar masses (typically $\lesssim 0.1$). The neutron star thus acts as a tugboat: it pulls an object more massive than itself, and thus its own velocity is diminished. When these neutron stars are eventually

Table 1 Binary kinematic data

	Star	D (kpc)	z (pc)	V_t^* (km s^{-1})	$V_r^\dagger$ (km s^{-1})	P (d)	Companion	Refs‡
I	HD 36486	0.3	90	5	+16	5.7	B2 V	2
	HD 57060	0.9	80	30	—11	4.4	Equal mass companion	2
	HD 163181	0.8	55	20	—40	12.0	Possibly O9	3
	HD 166937	0.15	5	4	—6	180.0	Eclipsing	2, 4
II	HD 108	1.0	30	110	—70	4.6	Low mass, invisible	2, 5
	HD 148937	0.8	20	(80)	—	—	Low mass, invisible	2, 6
	HD 152667	1.3	30	75	—35	7.8	Low mass, invisible	2
III	HD 77581	1.3	140	(100)	+20	8.95	Pulsating X-ray source	6, 7
	(3U 0900-40)							
	HD 153919	1.2	65	(60)	—65	3.4	Pulsating X-ray source	6, 7
	(3U 1700-37)							
	HD 226868	2.0	140	(280)	—2	5.6	High mass X-ray source	6, 7
	(Cyg X-1)							

*Transverse velocities in parentheses are only 1.5 – 2.0σ PM results from ref. 6.

†Radial velocities for OB supergiants may be strongly affected by photospheric expansion.

‡All information not in the cited references is from ref. 1.

released at the time of the second supernova, they will already be several Myr old. While passing through the X-ray binary stage they will have undergone a spin-down due to magnetospheric braking in the stellar wind of the companion and a subsequent spin-up to a period of $\lesssim 1$ s due to accretion torques (compare the evolutionary model for such massive systems of van den Heuvel¹⁹). Thus it is not surprising that their characteristic and dynamical ages disagree. They will be relatively low-velocity objects, with final space velocities composed of their former systemic and orbital velocities. The high orbital velocities observed for some of the neutron stars in X-ray binaries are apparently reduced by mass loss from the system before disruption. The system will apparently be disrupted when the second supernova occurs (the relative infrequency of binary pulsars attests to the high probability of disruption) and the second remnant neutron star will be accelerated by its asymmetric radiation force to a high velocity and join the Class B sources.

In this picture, the 10^3 -yr-old Crab pulsar at $z = 200$ pc and $v_t > 100$ km s⁻¹ is a result of the second supernova that disrupted the system which was carried to this height by the acceleration of the remnant from the earlier explosion. The binary pulsar system also apparently reached its present $z = 230$ pc in a similar manner, but the evolution of its companion has failed to disrupt the system. Many of the OB runaway stars which have typical kinematical ages of 1–5 Myr (ref. 20) may be binary systems that have been accelerated by the primary remnant, but have yet to reach the second supernova stage. The high velocity group 2 sources in Table 1 are a subset of these stars for which the existence of a low-mass companion has been established. They will turn on as X-ray sources a few Myr (ref. 19) later at higher z (group 3).

The fact that even the youngest pulsars and a large fraction²¹ ($\sim 30\%$) of the radio supernova remnants are found at $z > 100$ pc is consistent with a mechanism that drags some OB stars away from their birthplaces before they become supernovae.

The relative frequencies of OB binary systems, high latitude SNRs²¹, and Class A pulsars¹³ are all 20–30%; in addition, the total galactic population of Class A pulsars and systems able to evolve through the X-ray binary stage in $\sim 10^7$ yr are also in reasonable agreement. Future observations of more pulsar proper motions and searches for pulsars, X-ray sources and/or hidden, low-mass objects as companions to OB stars could better define this evolutionary picture.

DAVID J. HELFAND
EUGENE TADEMARU

Department of Physics and Astronomy,
University of Massachusetts,
Amherst, Massachusetts 01003

Received 17 January; accepted 7 March 1977.

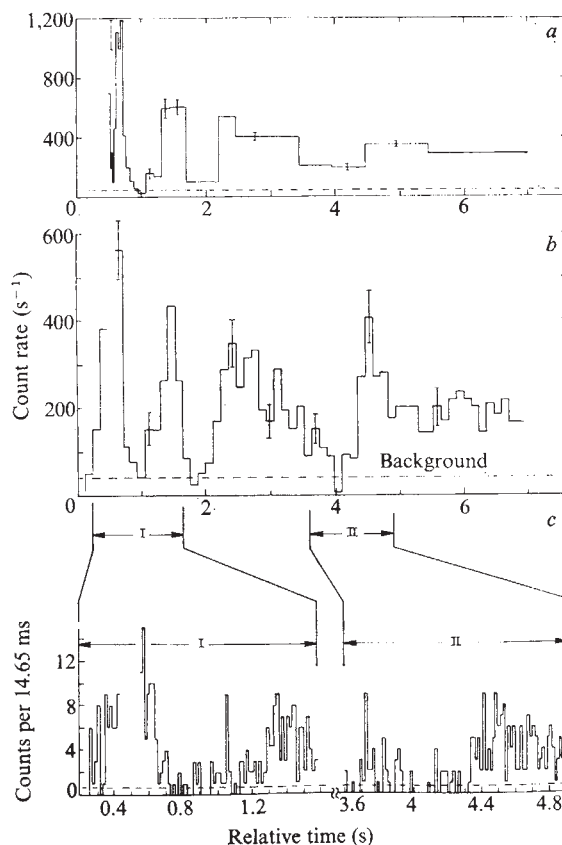
- ¹ Hutchings, J. B. *Astrophys. J.* **203**, 438 (1976).
- ² *Wilson General Catalogue of Stellar Radial Velocities* (Carnegie Institution of Washington, 1953).
- ³ Hutchings, J. B. *Publ. Astr. Soc. Pacific* **87**, 245 (1975).
- ⁴ *Catalogue of Bright Stars* (Yale University Observatory, 1964).
- ⁵ Hutchings, J. B. *Astrophys. J.* **200**, 127 (1975).
- ⁶ *Smithsonian Astrophysical Observatory Star Catalogue* (1966).
- ⁷ X-ray Binaries *Symp. Proc. NASA-GSFC* (National Aeronautics and Space Administration, Washington D.C., 1975).
- ⁸ Taylor, J. H. & Manchester, R. N. *Astrophys. J.* (in the press).
- ⁹ Manchester, R. N., Taylor, J. H. & Van, Y. Y. *Astrophys. J. Lett.* **189**, L119 (1974).
- ¹⁰ Anderson, B., Lyne, A. G. & Peckham, R. J. *Nature* **258**, 215 (1975).
- ¹¹ Backer, D. C. & Sramek, R. A. *Astr. J.* **81**, 430 (1976).
- ¹² Helfand, D. J., Taylor, J. H. & Manchester, R. N. *Astrophys. J. Lett.*, **213**, L1 (1977).
- ¹³ Helfand, D. J. & Tademaru, E. *Astrophys. J.* (in the press).
- ¹⁴ Gunn, J. E. & Ostriker, J. P. *Astrophys. J.* **160**, 979 (1970).
- ¹⁵ Lyne, A. G., Ritchings, R. T. & Smith, F. G. *Mon. Not. R. Astr. Soc.* **171**, 579 (1975).
- ¹⁶ Jones, P. B. *Astrophys. J.* **209**, 602 (1976).
- ¹⁷ Harrison, E. R. & Tademaru, E. *Astrophys. J.* **201**, 447 (1975).
- ¹⁸ Paczynski, B. *Ann. Rev. Astr. Astrophys.* **9**, 183 (1971).
- ¹⁹ van den Heuvel, E. P. J. *Proc. Eighth Texas Symp. Relativistic Astrophys.* (in the press).
- ²⁰ Blaauw, A. *Bull. astr. Insts Netherlands* **15**, 265 (1961).
- ²¹ Tamman, G. *Proc. Eighth Texas Symp. Relativistic Astrophys.* (in the press).

Preliminary results from Solrad 11 γ -burst detectors

We present here temporal and 0.2–2 MeV spectral data from two γ bursts observed on 12 June and 16 August 1976, by detectors on the Solrad 11A and 11B satellites. The 12 June burst showed evidence for structure on time scales down to ~ 10 ms throughout its lifetime, whereas the 16 August burst varied only with characteristic times longer than a few tenths of a second. A search for both slow and fast spectral variability gave negative results. Accurate absolute burst times are, however, not yet available, but since both bursts were also observed by at least one Vela satellite, positions are calculable and will be reported.

The Solrad 11A and 11B spacecraft were launched in March 1976, and are currently positioned on opposite sides of the Earth at heights of about 20 Earth radii. The γ -burst instrument on each spacecraft is omnidirectionally sensitive over the 0.2–2 MeV range and consists of a pair of $1\frac{1}{2}$ inch by $1\frac{1}{2}$ inch CsI scintillation detectors, the combined output of which drives circuitry designed for response only to rapid, transient events. A total average count rate during any 0.625 s interval of greater than about 7 sigma above background triggers the system, causing temporal and spectral information—including 0.6 s of pre-trigger data—to be stored in a memory for later telemetering to Earth. The temporal information is stored as data words which can be either of two types. For count rates below $1,092$ s⁻¹ the total number of counts accumulated per 14.65 ms is

Fig. 1 Count rate plotted against time for the 12 June burst. The time origin of the graphs is at $1903:40$ UT ± 2 s. *a*, Response of Vela 5A, which triggered at about the same time as Solrad 11A, but which does not store pre-trigger data. The full time resolution capability of this and the other Vela detectors is displayed here. *b*, Eight-channel (117.2 ms) time averages of Solrad 11A data for the same burst. *c*, Segments I and II of the Solrad data, but with 14.65 ms resolution.



recorded. At higher rates the time interval required to accumulate 16 counts is recorded with 0.3 ms resolution. Counts are simultaneously accumulated into four spectral channels covering the energy ranges of approximately 0.2–0.3, 0.3–0.4, 0.4–0.6, and 0.6–2.0 MeV. The contents of the spectral channels are read into the memory after each eight time accumulations, so the time resolution for spectral changes can range from 2.4 to 117.2 ms. Since there is no provision for in-flight energy calibration and the spacecraft mass complicates the detector response, the spectral data can be used to search for spectral variability, but not to obtain reliable spectral shapes.

Figure 1 shows count rate plotted against time for the burst of 12 June. Since the detector high voltage on Solrad 11B was turned off at the time of this burst, the only data available are from Solrad 11A and the Vela satellites. We have used two different time scales to show the overall burst structure and to give a higher time resolution view of selected portions of the burst. The 12 June burst was relatively intense, and was observed by γ -burst detectors on four Vela satellites. Figure 2 is a similar presentation of the 16 August burst. The response of 11A is low relative to 11B due to the probable malfunction of one of its two CsI detectors. This burst is small by Vela standards, having been just intense enough to trigger one Vela detector.

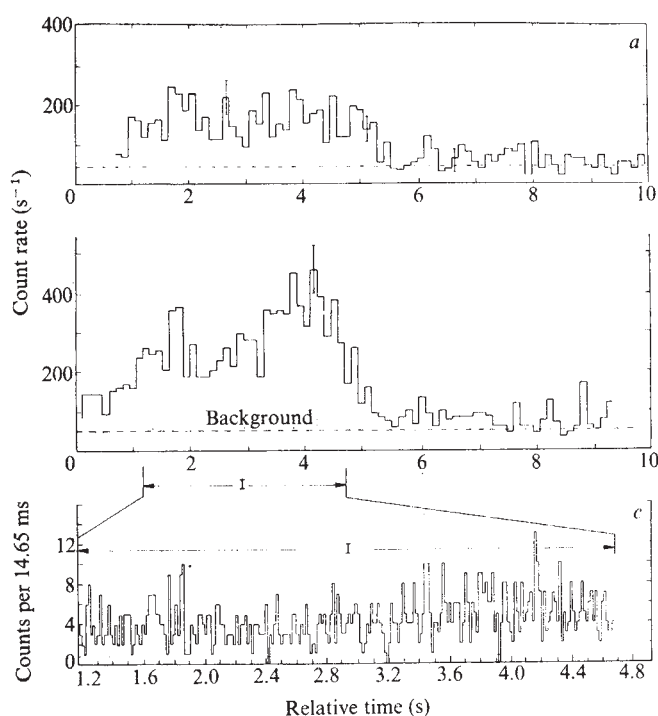


Fig. 2 Count rate plotted against time for the 16 August burst. The time origin is at 1615:31 UT $\pm$ 2 s. *a* and *b*, 117.2-ms averages of the responses of Solrad 11A and 11B, respectively. *c*, Section I of the Solrad 11B data replotted with 14.65-ms resolution.

The high time resolution graphs (Figs 1c and 2c) demonstrate an important difference in the signatures of the two bursts. The 12 June burst is very 'spiky', and shows clear evidence for structure on time scales as fast as ~ 10 ms. The fact that finer structure is not seen might simply be due to insufficient detector sensitivity, since the time resolution becomes better than 14.65 ms only for count rates higher than $1,092 \text{ s}^{-1}$. The 16 August burst, on the other hand, shows statistically significant structure only on time scales slower than a few tenths of a second. Since the two bursts were detected at about the same statistical level, the apparent difference between them is almost certainly real.

We have not explicitly shown rates in any of the separate spectral channels, but we have examined these data for evidence of spectral variability. Specifically, we carried out a χ^2 test of

the assumption that the hardness ratio—that is, the ratio of the flux in channels 3 and 4 to the flux in channels 1 and 2—remained constant throughout each burst. That is, we examined the quantity

$$\chi^2 = \sum_i \left(\frac{r_i - R}{\sigma_i} \right)^2$$

where r_i is the hardness ratio in a given spectral accumulation period, σ_i is an estimate of the standard deviation of r_i , and R is a free parameter that is the hardness ratio averaged over the burst. In several instances the probability distribution functions of the r_i differed significantly from the normal probability distribution, thus complicating the χ^2 analysis. (Note that a random variable that is the ratio of even normally distributed random variables is not approximately normally distributed unless the denominator in the ratio has a small fractional standard deviation.) Although we accounted for this effect in the worst cases—that is, r_i calculated from the smallest statistical samples—the χ^2 values obtained should be considered as only approximate indicators of goodness-of-fit.

To carry out the most sensitive possible search for both slow and fast variability, we carried out various time averages of the data before applying the χ^2 test. Basically, we found no compelling evidence for spectral variability, although for some of the time averages in the 0.25–1 s range χ^2 probabilities were below $\sim 10\%$. Since the spectral variability suspected on the basis of these probabilities was not correlated with the overall temporal structure of the bursts, however, we felt that the low probabilities were due partly to chance and partly to the statistical complications discussed above. The uncertainty in R for either burst was 5%, implying spectral constancy to within $\sim 10\%$ on several-second time scales.

In summary, we have seen evidence in one of two γ bursts for temporal structure on time scales down to ~ 10 ms. Higher detector sensitivity might have revealed faster time structure. We have also seen evidence that over the 0.2–2 MeV range the spectra of the two bursts were constant on several-second time scales to within $\sim 10\%$ during the observed—that is, most intense—portions of the bursts. This evidence favours theoretical burst models having the following features: The emission region can be at least as small as a white dwarf. The temporal smoothness of the 16 August burst suggests, however, that some burst sources might be much larger in at least one dimension, or that the γ rays sometimes undergo considerable dispersion—for example, by Compton scattering—between the emission region and the Earth. The source of the burst energy is not much affected by the burst itself, nor does it seem that the source is undergoing rapid physical change during the course of the burst. That is, we do not see evidence for physical processes such as rapid cooling or magnetic field changes that would be expected to cause spectral variations.

We thank Don Horan and Andy Fox of NRL and the crew at the Blossom Point tracking station for enabling us to obtain data in spite of formidable software problems. This research was supported by the US Energy Research and Development Administration.

J. G. LAROS
W. D. EVANS
R. W. KLEBESADEL
R. A. OLSON

University of California,
Los Alamos Scientific Laboratory,
Los Alamos, New Mexico 87545

Sandia Laboratories,
Albuquerque, New Mexico 87115

R. E. SPALDING

Received 8 February; accepted 7 March 1977.

Detection of ^4He in stratospheric particles gives evidence of extraterrestrial origin

WHIPPLE predicted the existence of micrometeorites, small interplanetary particles which enter the Earth's atmosphere without being melted by frictional heating¹. During the past 2 yr, large numbers of micrometre sized stratospheric particles have been collected which are believed to be extraterrestrial, because their elemental compositions closely match those of primitive meteorites^{2,3}. We report here the detection of large concentrations of ^4He in some of the particles. This not only suggests an exposure to solar wind, but also indicates that these particles are true micrometeorites in the sense that they were not strongly heated by entry into the atmosphere. Strong heating would have caused much of the helium to escape.

The particles were collected by an inertial impaction collector attached to a NASA U-2 aircraft flown at 20 km. The collector consisted of a 20-cm² oil-coated collection surface which was rammed through ambient air at the aircraft's cruise speed of 200 m s⁻¹. With the exception of aluminium oxide, a common stratospheric contaminant in this size range that is thought to come from solid fuel rocket propellant⁴, more than half the particles collected are believed to be of extraterrestrial origin. Nondestructive X-ray analyses have shown that most of these particles have abundances of Fe, Mg, Si, C, S, Ca, Al, Ni, Cr and Mn which closely match those of type C-1 carbonaceous chondrite meteorites^{2,3}. A total of 250 particles with diameters $\geq 4\text{ }\mu\text{m}$ have been collected by sampling about $3 \times 10^5\text{ m}^3$ of stratospheric air.

Surfaces exposed to the interplanetary medium usually contain relatively large amounts of ^4He due to implantation of solar wind⁵. Ions from solar wind implant to depths of $\sim 500\text{ }\text{\AA}$ in solid surfaces and form helium-rich layers on all exposed areas. Typical lunar soils have solar wind contents of the order of $0.1\text{ cm}^3\text{ (STP) g}^{-1}$. Lunar soil breccias and gas-rich meteorites also contain solar wind helium, but at generally lower levels due to diffusive losses and other processes. The flux of ^4He from the solar wind, as measured from the Apollo solar wind foil experiments⁶, varies from 4×10^6 to $18 \times 10^6\text{ cm}^{-2}\text{ s}^{-1}$. From these measurements, it can be calculated that for $10\text{-}\mu\text{m}$ -sized grains exposed to solar wind at 1 AU, a saturation level

of helium would be reached on a time scale of 10–100 yr. The level of saturation depends on the grain temperature and mineralogy, but is expected to be in the range 0.01 to $1\text{ cm}^3\text{ (STP) g}^{-1}$.

To verify the extraterrestrial origin of the collected stratospheric particles, ten were analysed for helium using a high-sensitivity gas mass spectrometer. Seven of the ten provided measurable quantities of ^4He (Table 1). Although in several cases the helium concentration levels were quite high, the absolute amounts of helium measured were always small, of the order of 10^9 – 10^{10} atoms, which made it necessary to use the specialised mass spectrometer system and special techniques.

The stratospheric particles were prepared for helium analysis by mounting them individually in small aluminium foil sandwiches. Particles were transferred to the foils under a stereo microscope using a hand-held $8\text{-}\mu\text{m}$ diameter glass fibre. Silicone oil was used in the transfer and $\sim 10^{-10}\text{ g}$ of residual oil was retained on each particle. After particle transfer, each foil was folded to form a sandwich to enclose and protect the particle. The mass of aluminium ranged from 2×10^{-3} to $7 \times 10^{-3}\text{ g}$. In addition to the ten foil sandwiches which contained particles, fourteen sandwiches containing only the silicone oil drops were also prepared for use as control samples. For each set of mass spectrometer runs, the masses of the aluminium control foils closely matched the masses of the sample foils.

The mass spectrometer system we used⁶ involves no new concepts, but has refinements over other systems^{7–9}, giving simplicity of operation and an ability to detect low concentrations of helium in small specimens. Helium content was determined by melting and vaporising each aluminium foil packet in a resistance-heated tungsten coil under a vacuum. The released helium ($\sim 10^9$ atoms ^4He) was quickly passed over liquid nitrogen-cooled activated charcoal and non-vaporisable Al–Ti getters operated at room temperature⁶. Two minutes after vaporisation, a known fraction of the released helium gas was allowed directly into the mass spectrometer volume which had just been isolated from its vacuum pumps for 'static mode' operation. Absolute calibration of the instrument sensitivity was accomplished by adding a precisely known number of atoms of ^3He ($\sim 1 \times 10^{13}$ atoms) to the sample vacuum enclosure, and following the identical procedure to introduce the gas to the mass spectrometer. These low level ^3He spikes, which

Table 1 Helium measurements in stratospheric particles

Sample No.	Estimated mass (ng)	Released by particle, plus system background	Measured amount of ^4He and 1σ uncertainty (in units of 10^9 atoms)		Net ^4He released by particle	Concentration of ^4He in particle ($\text{cm}^3\text{ (STP) g}^{-1}$)
			Released by control, plus system background	Range of system background		
SP-2	0.34	5.7 ± 0.2	3.4 ± 0.2	3.2–3.5	2.3 ± 0.3	0.25 ± 0.03
SP-3	0.43	3.8 ± 0.2	3.4 ± 0.2	3.2–3.5	0.4 ± 0.3	0.035 ± 0.026
SP-9	0.33	3.6 ± 0.1	2.8 ± 0.2	2.7–3.0	0.8 ± 0.2	0.09 ± 0.02
SP-10	2.0	4.4 ± 0.1	2.8 ± 0.2	2.7–3.0	1.6 ± 0.2	0.030 ± 0.004
SP-11	0.55	3.4 ± 0.2	3.5 ± 0.2	3.0–3.6	–0.1 ± 0.3	≤ 0.02
SP-12	1.0	3.2 ± 0.2	3.5 ± 0.2	3.0–3.6	–0.3 ± 0.3	≤ 0.01
SP-14	6.8	30.2 ± 0.2	2.9 ± 0.1	2.7–3.0	27.3 ± 0.2	0.149 ± 0.001
SP-15	24	3.3 ± 0.2	2.0 ± 0.2	1.8–2.2	1.3 ± 0.3	0.0020 ± 0.0003
SP-16	0.79	2.0 ± 0.2	2.0 ± 0.2	1.8–2.2	0.0 ± 0.3	≤ 0.01
SP-17	0.21	3.2 ± 0.1	1.9 ± 0.1	1.7–2.0	1.3 ± 0.2	0.23 ± 0.04

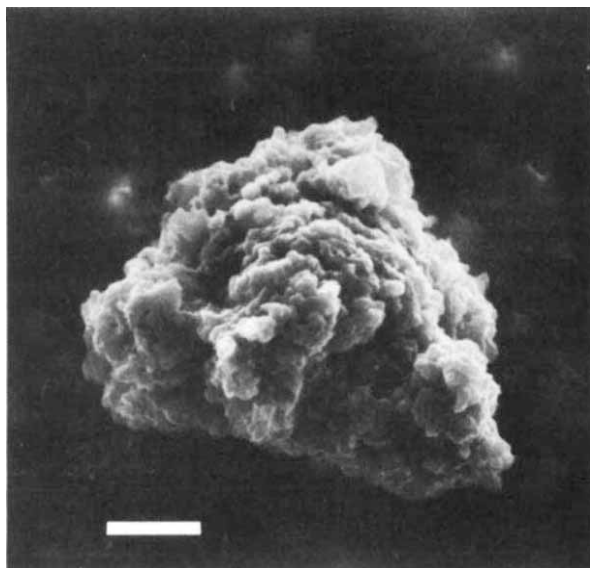


Fig. 1 Scanning electron micrograph of stratospheric micrometeorite SP-14. The elemental abundances of the particle match type 1 carbonaceous chondrite meteorite values within a factor of two, and as revealed at higher magnification, the particle is an aggregate of $\sim 1,000$ -Å grains. The scale bar represents 5 μm .

were obtained by partitioning through calibrated volumes, were preferred to ^4He because of possible 'memory' effects that the latter isotope might contribute. The usually preferred 'spiking' procedure of immediately mixing the ^3He with the unknown ^4He gas sample was not followed because the ^3He spikes also contained $\sim 1 \times 10^{10}$ atoms of ^4He . A nearly constant mass discrimination of the mass spectrometer was repeatedly measured during each series of runs by analysing precisely known mixtures of ^3He and ^4He introduced from other vacuum enclosures.

The results of helium analyses of samples and controls are given in Table 1. The measurements were complicated slightly by a nearly constant small residual helium background displayed by the mass spectrometer system, resulting from a steady permeation of helium through glass components and from the desorption of helium from previous larger gas samples. Although this background changed slightly from day to day, it varied little within any given day as shown in column 5 of Table 1. The uncertainty in the measurement of this background ($\sim \pm 2 \times 10^8$ atoms ^4He) was the major source of uncertainty in the results.

The sample vacuum enclosure has mountings for four tungsten coils so that a combination of two samples and two controls could be vaporised in 2 h. Comparison of the data from the control samples (column 4) with background values (column 5) shows that no perceptible additional helium was released either by the control samples or by the heating of the tungsten coils. The results in column 3 for seven of the ten stratospheric particles, however, are clearly outside the range of the system background. The net ^4He released by these particles (column 6) ranged from a barely perceptible 4×10^8 atoms for sample SP-3 to a sizeable 2.7×10^{10} atoms for SP-14.

Measurements made of the ^3He released by the particles and controls, not listed in Table 1, were subject to the memory effects of the ^3He spikes. Consequently, net measured ^3He values of 1.9×10^9 and 4×10^8 atoms, respectively from samples SP-2 and SP-15 are not considered to be real. The other samples showed no measurable ^3He above background levels.

The final column in Table 1 lists helium concentrations ranging from 0.002 to 0.25 cm^3 (STP) g^{-1} . Such high concentrations confirm that the particles were extraterrestrial,

and that some or all of them were exposed to solar wind for at least 10–100 yr. Since solar wind ions are implanted only to depths of ~ 500 Å, the measurements also indicate that the particles existed as small particles in space and were not produced in the atmosphere by fragmentation of larger meteoroids.

The possibility that the observed helium could instead have been the product of decaying uranium seems remote. Production of the high level of helium found in sample SP-14, for example, would require a uranium concentration of 200 weight parts per million (w.p.p.m.) and a decay time of 4.6×10^9 yr. Although phosphate and zircons, sometimes found as inclusions in meteorites, can contain up to ~ 20 w.p.p.m. uranium, non-destructive X-ray analyses of these particles have shown their compositions closely to resemble chondritic meteorites which generally have uranium concentrations three orders of magnitude lower than the latter figure.

The compositional similarity of all ten particles, together with the observed presence of helium in seven of them suggests that all ten are of extraterrestrial origin. The absence of observable helium in three of the particles and the widely varying concentrations observed in the others can be ascribed to varying amounts of heating during entry into the atmosphere. It is estimated that with typical entry parameters a 10- μm particle will experience a thermal pulse of 500 °C or more and that this pulse will last for several seconds. The particles in this study which showed high helium concentrations after entry into the atmosphere are true micrometeorites in the sense that they have not experienced any substantial heating. Micrometeorites thus join meteorites and lunar samples as sources of extraterrestrial materials which can be analysed in the laboratory for clues to early Solar System processes. Because micrometeorites have probably cometary origin¹⁰ they are potentially a valuable source of primitive Solar System matter.

R. S. RAJAN

*Department of Terrestrial Magnetism,
Carnegie Institution of Washington,
Washington DC 20015*

D. E. BROWNLEE

D. TOMANDL

P. W. HODGE

*University of Washington,
Seattle, Washington 98195*

HARRY FARRAR IV

R. A. BRITTEN

*Rockwell International,
Atomics International Division,
Canoga Park, California 91304*

Received 11 February; accepted 7 March 1977.

- Whipple, F. L. *Proc. natn. Acad. Sci. U.S.A.* 37, 19–30 (1951).
- Brownlee, D. E., Hörz, F., Tomandl, D. A. & Hodge, P. W. in *The Study of Comets*, NASA SP-393, 962–982 (National Aeronautics and Space Administration, Washington DC, 1976).
- Brownlee, D. E., Tomandl, D. A. & Hodge, P. W. in *Interplanetary Dust and Zodiacal Light* (ed. Elsasser, H. & Fechtig, H.) 279–283 (Springer, Heidelberg, 1976).
- Brownlee, D. E., Ferry, G. V. & Tomandl, D. A. *Science* 191, 1270–1271 (1976).
- Geiss, J., Buehler, F., Cerruti, H., Eberhardt, P. & Filleur, C. in *Apollo 16 Preliminary Science Report*, NASA SP-315, 14, 1–10 (National Aeronautics and Space Administration, Washington DC, 1972).
- Farrar, IV, H., McElroy, W. N. & Lippincott, E. P. *Nucl. Technol.* 25, 305–329 (1975).
- Schaeffer, O. A. & Zähringer, J. *Phys. Rev.* 113, 674–678 (1959).
- Signer, P. & Nier, A. O. *J. geophys. Res.* 65, 2947–2964 (1960).
- Alimova, I. A., Boltenev, B. S., Gartmanov, V. N., Mamyurin, B. A. & Shustrov, B. N. *Geokhimiya* 9, 1044–1049 (1966).
- Millman, P. M. in *From Plasma to Planet* (ed. Elviu, A.) 157–168 (Wiley, New York, 1972).

Pulsating aurora induced by upper atmospheric barium releases

AN effect which appeared to be pulsating aurora was observed to be associated with explosive releases of barium vapour near 250 km altitude, but only when the explosions were in the path of precipitating electrons associated with the visible aurora.

The primary purpose of this series of rocket experiments, launched from Poker Flat, Alaska ($65^{\circ} 77.5' \text{ N lat.}, 147^{\circ} 28.91' \text{ W long.}$) during March 1976, was the measurement of the magnitude and direction of electric fields and neutral winds in the auroral atmosphere¹. Each explosive charge was a standard 1.5 kg thermite mixture of Ba and CuO with an excess of Ba metal which was vaporised and dispersed by the thermite explosion². Traces of Sr, Na and Li were added to some of the charges in varying amounts, and an extensive set of spectrophotometric observations from the ground was planned to determine the effectiveness of each of these luminous, un-ionised species in tracing the upper atmospheric, neutral wind.

The two flights of interest here were made in the morning twilight on 27 and 28 March 1976, at about 1250 and 1300 UT, respectively. The approximate trajectory of the rockets is shown in Fig. 1 along with the location of the four releases and the ground stations. One difference between the two flights is that perigee on the 27 March flight was approximately 50 km higher than that for the 28 March flight.

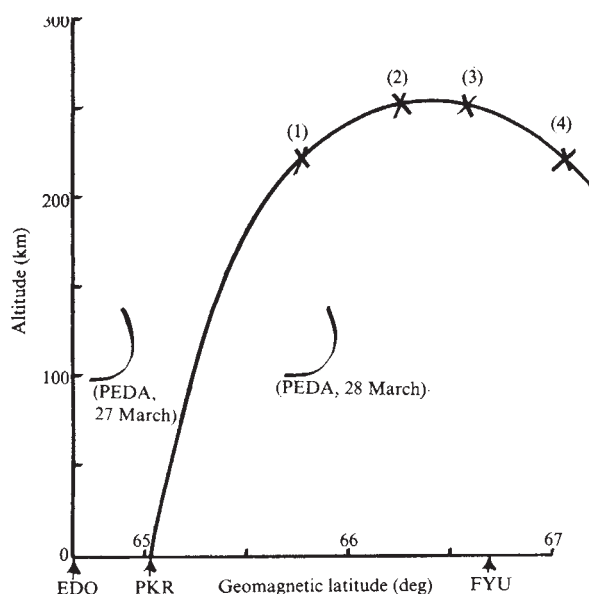


Fig. 1 An elevation view of the rocket trajectory on 28 March 1976, in the geomagnetic meridian ($\sim 30^{\circ} \text{E}$ of true N) through the Poker Flat Research Range. Shown also are the approximate positions of the observing sites (Ester Dome: EDO, Poker Flat: PKR, Fort Yukon: FYU), the barium releases and the natural aurora. Figures in parentheses are release nos. PEDA, poleward edge of diffuse aurora.

Spectrophotometric observations were made from two locations, Ester Dome ($64^{\circ} 52.8' \text{ N lat.}, 148^{\circ} 3.4' \text{ W long.}$) and Poker Flat. Emission spectra were obtained for all the releases using a television spectrograph at Ester Dome. With this instrument, spectra between 4,000 and 8,000 Å at 40 Å resolution of chemical release clouds as well as intense aurorae are recorded on video tape at a rate of 30 frames per second. Focusing the sky on the entrance slit gives an angular spatial coverage of 8° (resolution: $< 0.5^{\circ}$)

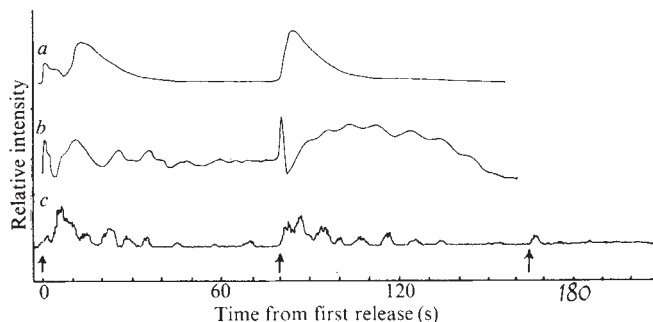


Fig. 2 Observations of two atmospheric emissions and one neutral Ba emission associated with upper atmospheric Ba releases made on 28 March 1976, from Poker Flat, Alaska. The times of three of the four releases are indicated by arrows. The fourth release showed neither enhancement nor pulsation of the atmospheric emission. a, 6,110 Å BaI, TV spectrograph; b, 5,577 Å OI, TV spectrograph; c, 4,278 Å N_2^+ , photometer.

along the spectral line. At Poker Flat an optical interference filter photometer with an 8° field-of-view was used to observe 4,278 Å (N_2^+ ING) emission in each release and record the amount of ionisation of atmospheric constituents. Numerous other instruments making support observations included two wavelength-scanning Ebert-Fastie spectrometers at Poker Flat, the Chatanika Incoherent Scatter Radar, and all-sky cameras, riometers and meridian scanning photometers at Poker Flat and Fort Yukon ($63^{\circ} 34' \text{ N lat.}, 145^{\circ} 17' \text{ W long.}$).

The morphology of the aurora on the occasion of both rocket flights was quite similar and characteristic of early morning aurorae in the Alaskan sector. Towards the equator was a large area of pulsating aurora and auroral glow. This area was terminated on the poleward side by a series of closely-spaced arcs appearing as striations in an E-W direction. Arc segments peeled off this boundary towards the pole with an overall eastward motion, forming so-called 'omega bands'. All of these features are part of what is referred to as the 'diffuse' aurora because of its apparent lack of structure to satellite-borne cameras³. The poleward, 'discrete' aurora consisted of weak, rayed arcs and arc segments. The location of the boundary between the diffuse and discrete auroral regions seemed to be relevant to the occurrence of the pulsating aurora effect which is the subject of this report. This location is shown schematically in Fig. 1.

On 28 March 1976, a distinct cloud of 5,577 Å [OI] emission was detected on the TV spectrograph located just below the bright Ba cloud on releases no. 1 and 2, but not on nos 3 and 4. The photometer operated at Poker Flat observed 4,278 Å (N_2^+ ING) emission associated with each release as shown in Fig. 2, together with the 5,577 Å data from the TV spectrograph. It is apparent that there is an enhancement of both emissions associated with the first two bursts and that these emissions pulsate with roughly a 10 s period for approximately 60–100 s following the burst. The 4,278 Å pulsations were of the order of 100 rayleighs against a background of 1 kR. This would be produced by a flux of approximately $5 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ of 500 eV electrons which would be deposited near the observed altitude of 200 km. Also from theory⁴, the corresponding intensity of 5,577 Å OI emission should be approximately 1 kR which is in reasonable agreement with TV spectrograph data, although that instrument is not calibrated on an absolute scale. No 5,577 Å OI emission was detected on March 1976, when all the clouds were poleward of the diffuse auroral boundary. Thus, because the enhancements occurred only when the chemical releases were above the diffuse aurora (or nearly so), it is tempting to assume that the presence of the diffuse aurora gives an indication of the conditions necessary to produce the pulsations.

Other geophysical effects associated with the early morning diffuse aurora were considerable absorption of galactic radio noise indicating D region ionisation produced by energetic auroral primary electrons, and electron densities in the E region $>10^5 \text{ cm}^{-3}$ observed by the Chatanika Radar. Table 1 summarises the observations from Poker Flat, indicating differences in the release environment which may be related to the production of the induced pulsations.

Table 1 Summary of observations during Ba releases

	27 March 1976 ~64.5° inv. lat.	28 March 1976 ~66° inv. lat.
Northern border of diffuse aurora		
D region radio wave absorption	Small	Large
E region electron density	$<10^4 \text{ cm}^{-3}$	$>10^5 \text{ cm}^{-3}$
F region electron density	$<10^4 \text{ cm}^{-3}$	$<10^4 \text{ cm}^{-3}$
Ionospheric electric fields	Weak, variable direction	Medium N-S directed, horizontal
Discrete aurora	Weak-rayed bands	None visible

The steady westward drift of ion clouds on 28 March indicated the existence of a N-S-directed, horizontal E field characteristic of early morning aurora. On 27 March the ion clouds striated and merged with the weak, discrete auroral bands indicating only variable, local E fields.

The TV spectrograph provides both spatial and spectral information on the development of the luminous cloud. The radius of the cloud was determined separately for the Li and Ba debris for release no. 3 on 27 March 1976 and both the Ba and Li clouds were observed to oscillate with a frequency near 1 Hz during the first 15 s after release. These high frequencies were not observed in the clouds which produced the auroral pulsations. Oscillations of longer periods (10 s) would be difficult to observe over the limited lifetime of the induced auroral pulsations. Thus, the observed high frequency oscillations may indicate the presence of lower frequency modes which could be intimately connected with the mechanisms producing the induced auroral pulsations.

Previous observations of enhanced auroral emissions and E region electron densities associated with F region Ba releases were made of the E region below the cloud by Stoffregen^{5,6} and although they showed the same overall time variation, there was no definite evidence of pulsations. In the case reported here, the diffuse aurora in the E region masked any enhancement which may have occurred and the observed pulsations were located just below the Ba cloud (~200 km altitude). Because the period of the pulsations was in the Alfvén or magnetoacoustic wave domain (10 s) and approximately that which is most often observed in the natural aurora at this latitude, attempts will be made in the near future to repeat the experiment in hopes of clarifying the mechanism for the natural production of pulsating aurora.

We acknowledge contributions to this work and its interpretation by J. Kan, D. Swift and W. Murcray. This work was supported by the US NSF and NASA.

C. DEEHR
G. ROMICK

Geophysical Institute,
University of Alaska,
Fairbanks, Alaska 99701

- ³ Lui, A. T. Y. & Anger, C. D. *Planet. Space Sci.* 21, 799 (1973).
⁴ Rees, M. H. & Luckey, D. J. *geophys. Res.* 79, 5181 (1974).
⁵ Stoffregen, W. J. *atmos. terr. Phys.* 32, 171-178 (1970).
⁶ Stoffregen, W. *Geophys. Publ.* 29, 515 (1972).

Configuration of the magnetic field and reconstruction of Pangaea in the Permian period

THE virtual geomagnetic poles of Laurasia and Gondwanaland in the Carboniferous and Permian periods diverge significantly when these continents are reassembled according to the fit calculated by Bullard *et al.*¹. Two interpretations have been offered: Briden *et al.*² explain these divergences by a magnetic field configuration very different from that of a geocentric axial dipole; Irving³ (and private communication), Van der Voo and French⁴ suggest a different reconstruction and it is shown here that these two interpretations are not incompatible and that the first may help the second.

I have developed a program which allows a spherical harmonic analysis of the palaeomagnetic field⁵ by a method similar to Kono's⁶. The magnetic field F_i at the point of colatitude θ_i and longitude ϕ_i is given by the equation

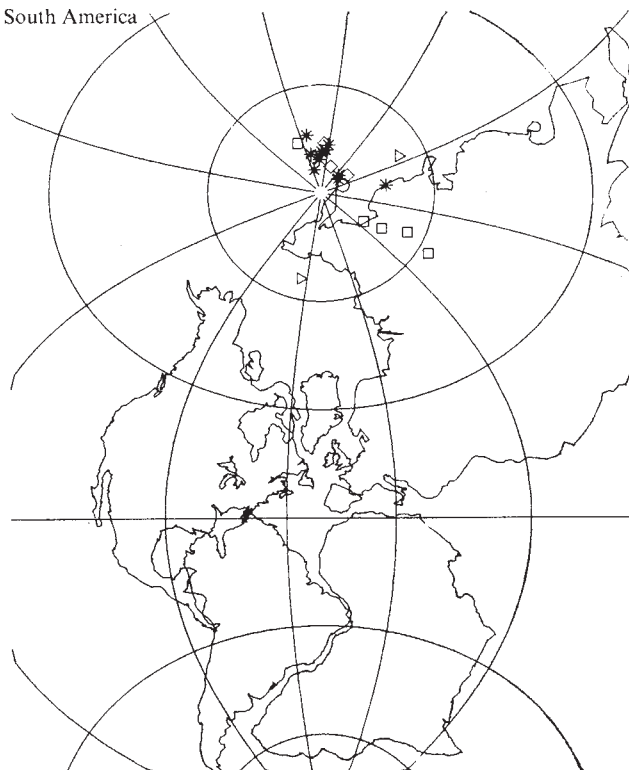
$$F_i = \sum G_j f_{ij}$$

where G_j is the expansion coefficient in spherical harmonics, and f_{ij} the vector whose components are trigonometric polynomial functions of θ_i and ϕ_i (notation used by Creer *et al.*⁷). A_i is the measured palaeomagnetic direction of unit length at the same point.

We need to minimise the following residual: $\Sigma(e_i)^2 = \Sigma(F_i/|F_i| - A_i)^2$ where $|F_i|$ is the modulus of F_i calculated starting from a first expansion. So a system of linear equations

Fig. 1 Proposed reconstruction. Transverse Mercator projection. The symbols indicate the location of the virtual geomagnetic poles that were used. Parallels and meridians are drawn every 30°.

- Morocco
- ▷ Africa
- * Europe
- ◇ North America
- South America



Received 10 December 1976; accepted 9 March 1977.

- ¹ Heppner, J. P., Stolarik, J. D. & Wescott, E. M. *J. geophys. Res.* 76, 6028 (1971).
² Föpl, H. *et al. Planet. Space Sci.* 13, 94-114 (1965).

with unknowns G_j is given, the system solved and new values calculated for $|F_i|$, and the calculation repeated until the solution is stabilised.

The first eight terms of the expansion allow calculation of the orientation of the dipole (geomagnetic poles) and the position of its centre (see formulae in Creer *et al.*⁷). I have applied this method to the problem of the magnetic field in the Permian and I have grouped the palaeomagnetic results for neighbouring sections and then analysed the data thus obtained separately for the two aggregates Africa–South America (8 groups) and Europe–North America (14 groups).

If these continents are reassembled according to Bullard's fit, with Africa in its present position, the results are as follows:

	geomagnetic axis		centre of the offset dipole	
			distance	direction
Laurasia	46.5°N	227°E	1170 km	33°S 205°E
Gondwana	29°N	238°E	780 km	29°S 183°E

In both cases the angle between the direction of the geomagnetic axis and the direction of the centre of the offset dipole is about 80°. This configuration is quite different from the mean axial field observed in the Tertiary, but is similar to the present-day field. This lateral offset of the dipole allows us to resolve the problem of relative longitudes of the two supercontinents.

I have first superposed both geomagnetic poles and then rotated one continent around this common axis until the directions in which the dipoles are offset are coincident. A relative rotation of Europe and North America of -35° around a pole situated at 22°N and 25°E (with respect to Africa) is a satisfactory solution and the corresponding reconstruction is shown in Fig. 1, with the common geomagnetic axis placed on the geographic pole. This reconstruction is close to that proposed by Irving³ for the same period, but with different premises.

M. WESTPHAL

*Institut de Physique du Globe,
5 rue Descartes,
F- 67084 Strasbourg, France*

Received 7 February; accepted 10 March 1977.

¹ Bullard, E. G., Everett, J. E. & Smith, A. G. *Phil. Trans. R. Soc. A* **251**, 41–51 (1965).

² Briden, J. C., Smith, A. G. & Sallomy, J. T. *Geophys. J. R. astr. Soc.* **23**, 101–117 (1970).

³ Irving, E. *Paleomagnetism and its Application to Geological and Geophysical Problems*, p. 270 (Wiley, New York, 1964).

⁴ Van der Voo, R. & French, R. B. *Earth Sci. Rev.* **10**, 99–119 (1974).

⁵ Westphal, M. *Thesis, Strasbourg* (1976).

⁶ Kono, M. *Rock magnetism and Paleogeophysics, Tokyo* **1**, 118–123 and 124–129 (1973).

⁷ Creer, K. M., Georgi, D. T. & Lowrie, W. *Geophys. J. R. astr. Soc.* **33**, 323–345 (1973).

Isotopic evidence for dramatic climatic changes in East Africa during the Pleistocene

RAINFALL decreased dramatically in the Lake Turkana region 1.8–2.0 Myr ago and in the Olduvai Gorge region 0.5–0.6 Myr ago. This is documented by a major increase in the $\delta^{18}\text{O}$ values of pedogenic and groundwater carbonates at these times. The data suggest that meteoric water of the earlier, more humid climate was 2–4 per mil lower in ^{18}O content than modern waters of these regions.

A terrestrial stratigraphic record covering a large portion of the Pleistocene is found in several sedimentary basins in the eastern Rift zone of East Africa. We have examined the stable isotope variations through the Pleistocene at Olduvai Gorge in northern Tanzania and in the sediments surrounding Lake Turkana in northern Kenya (Fig. 1). Thirty-three samples of carbonates were analysed as were several samples of modern rain and lake waters from the two areas. Oxygen isotope compositions of carbonates are controlled by the

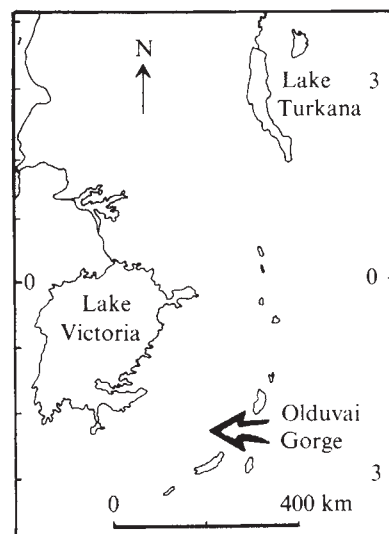
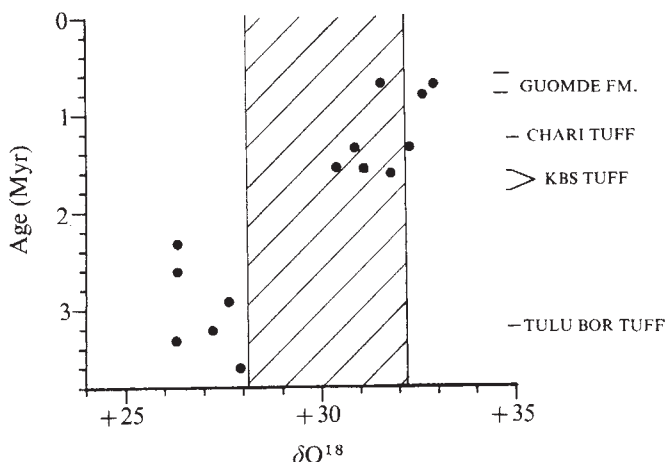


Fig. 1 Location map.

oxygen isotope compositions of the formation waters and by temperature¹. The oxygen isotope compositions of meteoric waters, in turn, are controlled by climatic factors^{2,3}. Consequently, isotopic analyses of groundwater carbonates provide excellent indicators of climatic conditions prevailing at their time of formation.

Three types of carbonates were analysed: pedogenic calcretes, pedogenic nodules, and groundwater nodules from lacustrine clays. The calcretes are impervious and represent sealed environments. This is documented by ^{14}C dating of 17 samples of upper Pleistocene and Holocene calcrete from Olduvai Gorge⁴. The dates fit with those determined on organic material from the same sequence and there is no overlap between the different stratigraphic levels. Further details of age relationships at Olduvai Gorge are documented by Hay⁴ and at Lake Turkana by Curtis *et al.*⁵ and Fitch and Miller⁶. Although some controversy still exists as to the age of the KBS Tuff at Lake Turkana, this tuff marks the climatic change discussed later. The Olduvai calcretes comprise both massive and laminated varieties, both of which were formed by meteoric water percolating downward through the soil column⁷. The

Fig. 2 A plot of $\delta^{18}\text{O}$ values of carbonates against age at Lake Turkana. The hatched area represents the isotopic compositions of theoretical carbonates in isotopic equilibrium with waters that have the range of $\delta^{18}\text{O}$ values for rain measured at Lake Turkana and Olduvai Gorge. The values were calculated for 25°C using the equation of O'Neil *et al.*⁸. Stratigraphical relations are shown at the right.



lacustrine nodules appear relatively impervious and presumably also represent sealed environments; also, groundwater nodules occur in relatively impermeable clays.

In the section at Lake Turkana, samples older than 1.8 Myr have $\delta^{18}\text{O}$ values in the range +26.3 to +27.9, whereas those younger are relatively enriched with a $\delta^{18}\text{O}$ range of +30.4 to +33.0 (Fig. 2). At Olduvai Gorge, the Masek Beds which formed 0.4 to 0.6 Myr ago⁴ mark a similar transition from light to heavy carbonates: Beds I, II, III and IV have $\delta^{18}\text{O}$ values of +24.1 to +27.1 and the Ndutu, Naisiusiu and Holocene sediments have $\delta^{18}\text{O}$ values of +30.9 to +31.8 (Fig. 3). At Olduvai, a dramatic event took place during Masek time and perhaps a minor one occurred at the end of Bed I (1.7 Myr ago⁴). A general increase in $\delta^{13}\text{C}$ with $\delta^{18}\text{O}$ was also observed in both areas.

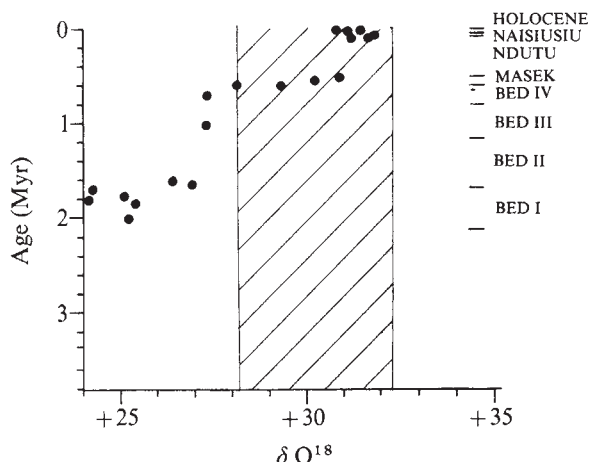


Fig. 3 As for Fig. 2, for Olduvai Gorge.

Craig⁹ has shown that East African waters are abnormally rich in the heavy isotopes of oxygen and hydrogen. The isotopic data do not fall on a line characteristic of worldwide meteoric water ($\delta D = 8\delta^{18}\text{O} + 10$) but are instead on a line characteristic of evaporating waters ($\delta D = 5\delta^{18}\text{O} + 10$). This is a result of the fact that the rain partially evaporates as it falls through the dry East African atmosphere. Analyses of five rain waters from Olduvai Gorge and Lake Turkana ranged in δD from +4.9 to +26.0 and in $\delta^{18}\text{O}$ from -0.4 to +3.5. The δD and $\delta^{18}\text{O}$ values of Lake Turkana and Lake Ndutu (west of Olduvai Gorge) are +37.1 and +5.84, and +41.7 and 6.7, respectively. These measurements concur with Craig's observations of unusually heavy waters in East Africa. Rain water has the lightest isotopic values of all the waters possibly involved in carbonate formation in this area. Modern waters from East Africa could not produce the relatively low $\delta^{18}\text{O}$ values observed in the older sediments unless soil and groundwater temperatures were 35–45 °C. This is completely unreasonable on the basis of faunal evidence. Therefore the soil and groundwaters from which the older carbonates precipitated originated from rainfall that was at least 2–4 per mil lower in ^{18}O content than found in modern waters of the region. This is favoured by lower evaporation rates and therefore implies that there was more rain in the past. Moreover, increased humidity might be expected to result in a decrease in temperature, which would cause a further decrease in $\delta^{18}\text{O}$ values of rainfall and groundwater. These results point to a drastic decrease in rainfall about 1.8–2.0 Myr ago at Lake Turkana and 0.5–0.6 Myr ago at Olduvai Gorge.

The Olduvai Beds contain faunal, lithologic and mineralogic evidence of increasing aridity from Bed I with a semi-arid climate, to the Masek Beds with an arid climate similar to that prevailing today. Bed I contains faunal

elements (e.g., the urocyclid slug) indicating moister conditions than any later deposit at Olduvai; the Masek Beds have forms characteristic of drier climates (e.g., the land snail *Limicolaria*)⁴. Extensive deposits of wind-worked tuffs occur from the Masek Beds to the present indicating drier conditions. The minerals formed by weathering in the Masek Beds (illite, zeolites, dawsonite) also indicate a climate drier than recorded previously at Olduvai Gorge⁴.

The sediments surrounding Lake Turkana also show evidence for a climatic change that coincides with the isotopic anomaly. Volcanic ash below the KBS Tuff (1.8 Myr old⁵) is altered to montmorillonite whereas that above it is altered to montmorillonite, chabazite and clinoptilolite. The shoreline of the lake that is older than 1.8 Myr was stable indicating that it probably was an open basin, whereas the fluctuating nature of the post-KBS shoreline and the presence of zeolites indicates that the lake periodically fell below overflow level.

The causes of the change in rainfall patterns are not clear from the study of these two areas. The climate of East Africa today is controlled by the intertropical convergence zone and the monsoons from the Indian Ocean^{10,11}. Any change in the intertropical convergence patterns, the relative positions and amount of highlands and their rain shadow effect, or in large bodies of water such as Lake Victoria, could have drastic effects on the amount of rainfall.

We are now determining the hydrogen and oxygen isotope compositions of hydrous authigenic minerals in an attempt to characterise more closely the isotopic composition of palaeometeoric waters in the region.

Field work was supported by the National Geographic Society, the National Science Foundation, the Penrose Fund, and the National Museum of Kenya. We thank L. D. White for his careful water analyses.

T. E. CERLING
R. L. HAY

Department of Geology and Geophysics,
University of California, Berkeley,
Berkeley, California 94720

J. R. O'NEIL

United States Geological Survey,
Menlo Park, California 94025

Received 14 February; accepted 7 March 1977.

- Epstein, S., Buchsbaum, R., Lowenstam, H. A. & Urey, H. C. *Bull. Am. Geol. Soc.* **64**, 1315 (1953).
- Epstein, S. & Mayeda, T. K. *Geochim. Cosmochim. Acta* **4**, 213 (1953).
- Dansgaard, W. *Tellus* **16**, 436 (1964).
- Hay, R. L. *Geology of Olduvai Gorge* (Univ. Calif. Press, 1976).
- Curtis, G. H., Drake, R. E., Cerling, T. E., Cerling, B. W. & Hampel, J. *Nature* **258**, 395 (1975).
- Fitch, F. J. & Miller, J. A. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y. et al.), 123 (Univ. Chicago Press, 1976).
- Gile, L. H., Peterson, F. F. & Grossman, R. B. *Soil Sci.* **101**, 347 (1966).
- O'Neil, J. R., Clayton, R. N. & Mayeda, T. K. *J. chem. Phys.* **51**, 5547 (1969).
- Craig, H. *Science* **133**, 1702 (1961).
- Flint, R. F. *Bull. Geol. Soc. Am.* **70**, 343 (1959).
- Griffiths, J. F. *Climates of Africa, World Surv. Climatology* **10** (1972).

Location and age of Mesolithic coastal occupation sites on Oronsay, Inner Hebrides

THE existence of at least five Mesolithic occupation sites on the island of Oronsay, Inner Hebrides, has now been established (Fig. 1). Preliminary results of recent excavation of two of the sites were reported in 1971 (ref. 1). Since then further excavations have been made and the locations and ages of the sites in relation to the position of the shore line of the sea at the maximum of the Holocene marine transgression have been investigated. The results obtained suggest that the ages of some of the sites in relation to the former shore line cannot be established solely on stratigraphical

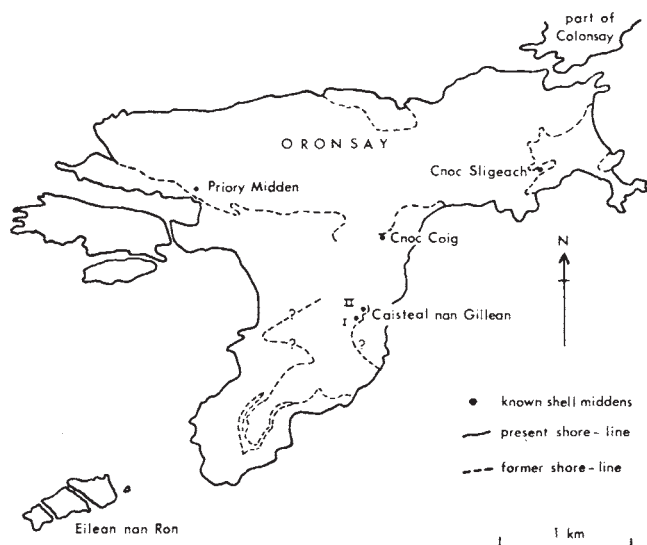


Fig. 1 Map of the Scottish Hebridean island of Oronsay (56° N, 6° W), showing the locations of known Mesolithic occupation sites and the position of the visible parts of the former shoreline mentioned in the text.

and altitudinal evidence but require supplementary information based on radiocarbon dating. It will be some time before the full details will be published, and so a short account of the results so far obtained is presented now.

The position of the former marine shore line differs markedly from that of the present coastline of the island of Oronsay (Fig. 1). The old shore line is obscured in places by blown sand, especially in the area between Caisteal nan Gillean and Cnoc Coig. It is possible, however, that at the maximum of the Holocene marine transgression Oronsay consisted of two main islands: a larger area of relatively high ground and a smaller low rocky area somewhat similar in aspect to the present Eilean nan Ron (Fig. 1) which supports a seal colony and is surrounded by limpet-inhabited rock skerries exposed at low tide. The islands may have been connected with each other at low tide as Oronsay is with Colonsay at present.

Measurement of high water level and low water level of spring tides in the summers of 1973, 1974, 1975 and 1976 at a number of stations along the eastern and western shores of Oronsay showed that present (summer) maximal tidal range varies between ~ 3.5 and 4.3 m, spring tidal range for practical purposes being 4.0 m, from -2.0 m to $+2.0$ m in relation to Ordnance Datum, Newlyn². If the tidal range around Oronsay at the time of the maximum of the Holocene marine transgression was approximately the same as at present, the contemporaneous Mean Tide Level (which is approximately the same as Mean Sea Level but is not precisely equal to it³) was $\sim +7$ m Ordnance Datum (OD) and the intertidal zone of that time was the area now standing at $\sim +5$ to $+9$ m OD. Measurements of the altitude of cobble debris deposited by modern severe storms (those of January and February 1974) suggests that along most of the coast of Oronsay similar storms at the maximum stand of the Holocene marine transgression may have thrown debris to locations that now occur up to $+11$ m OD. Storm debris might be expected to accumulate to somewhat higher levels than this on the crests of the cobble ridges that fringe the exposed southern peninsula of the former coast. The highest points on the crests of these ridges in fact occur at $\sim +11.4$ m OD (western ridge), $\sim +12.7$ m OD (southern ridge) and $\sim +13.1$ m OD (eastern ridge).

Three of the known Mesolithic occupation sites are located in positions that are above the highest probable high

water levels of ordinary spring tides and severe storms of the Holocene marine maximum. Occupation material of the Priory Midden occurs at levels between $\sim +12.8$ and $+14.3$ m OD. In the vicinity of the site the highest part of a (probably contemporaneous) marine cobble ridge occurs at $\sim +9.0$ m OD; other possibly contemporaneous beach deposits occur up to $\sim +11.9$ m OD. At Caisteal nan Gillean I occupation material is located on the crest of a sand mound between $\sim +16.0$ to $+17.3$ m OD and on the perimeter of the same mound at $\sim +11.1$ m OD. On the adjacent site of Caisteal nan Gillean II occupation material occurs at $\sim +12.5$ to $+15.5$ m OD, possibly fractionally lower, but certainly higher than $+11.0$ m OD, the probable upper limit reached by waters of severe storms when the Holocene sea stood at its highest level.

The two remaining known occupation sites, Cnoc Coig and Cnoc Sligeach, are located closer to the former shore line, both in altitude and position, than any of the other sites. Cnoc Coig is a low mound of (mainly) invertebrate shell debris, its crest at $\sim +10.2$ m OD, and the debris, up to ~ 1 m thick, rests on sand (possibly blown or beach sand) that in turn rests on solid rock. The shell mound is bordered

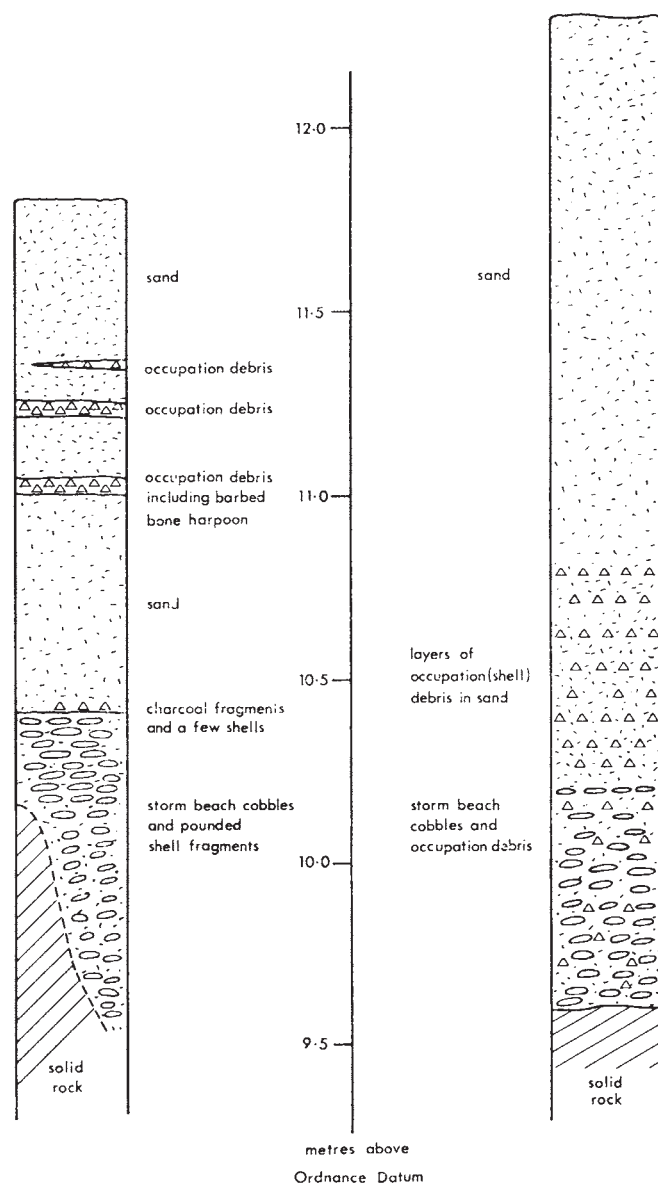


Fig. 2 Simplified vertical sections of recent excavations 30 m to the south of the crest of Cnoc Sligeach (left) and of excavations in 1913 by W. H. Bishop 20 m to the south-east of the crest of Cnoc Sligeach (right).

Table 1 Radiocarbon age determinations at Mesolithic occupation sites on Oronsay

Site	Nature of material dated	Position of sample within occupation debris	Lab no.	Age of sample in radiocarbon years (b.p.)
Caisteal nan Gilleann II	Wood charcoal { Humate abstract	Upper half of midden	Birm-346	4,920 ± 400
	Wood charcoal { After alkali pretreatment		Birm-347	5,150 ± 380
	Shells of <i>Patella vulgata</i>	Base of midden	Birm-348	5,450 ± 140
				5,570 ± 140 (outer)
Cnoc Coig		Base of midden		5,720 ± 140 (middle)
				5,850 ± 310 (inner)
	Wood charcoal	Base of midden	Q-1355	5,460 ± 65
	Wood charcoal	Upper part of midden	Q-1352	5,430 ± 130
	Wood charcoal { Two separate samples		Q-1351	5,495 ± 75
	Wood charcoal { From same horizon		Q-1354	5,535 ± 140
Cnoc Sligeach	Wood charcoal { Two separate samples	Base of midden	Q-1353	5,645 ± 80
	Wood charcoal { From same horizon		GX-1903	5,015 ± 210
	Valve of <i>Ostrea</i> sp.	Positions not known	GX-1904	5,755 ± 180
	Animal bone fragments (collagen)		BM-670	5,426 ± 159
	Wood charcoal	Upper part of midden (composite sample)		

on the north by a low rock exposure that was immediately adjacent to, if it did not itself form, the margin of the shore at the maximum of the Holocene marine transgression. The lowest level at which occupation material occurs at this site is ~ +9.0 m OD, approximately the same altitude as the estimated normal high water level on Oronsay at the maximum of the Holocene marine transgression. On this basis it is arguable that the Cnoc Coig site was occupied at the time of the marine maximum. Other evidence, however, argues strongly against this: the site is located on the seaward side of the rock exposure that (conjecturally) formed the contemporaneous shore line; fifty metres to the north of the site Holocene storm debris is present at levels up to ~ +9.5 m OD, but at the site itself no traces of storm debris in association with occupation material were found in the course of recent excavations. It may be concluded that the site certainly was not occupied during storms at the time of the marine maximum, and it is probable that the site was never occupied until after the sea had receded from its maximal position.

Occupation material at Cnoc Sligeach occurs near the crest of the mound at ~ +16.9 m OD, on the sides of the mound and around several centres located on the periphery of the mound. The lowest level at which substantial quantities of occupation material were found recently at this site was ~ +10.4 m OD, the material concerned (wood charcoal in association with shells of *Patella* sp. and *Littorina littorea*) occurring on the southern periphery of the mound and resting directly on storm-beach cobbles. In addition, a few small, diffuse, dark-brown or black patches, possibly of charcoal, were found sporadically distributed within the mixture of storm beach cobbles and pounded fragments of *Patella* sp. that extended downwards to solid rock at the same location (Fig. 2). The evidence is comparable with, but not precisely the same as, that described by Bishop¹ whose extensive excavations on the south-eastern flank of the mound in 1913 demonstrated that the maximum altitude at which marine cobbles were deposited there was ~ +10.2 m OD. Bishop also noted that, at the site of his investigation, 'throughout the entire depth of this beach were found the same implements and remains as were found in the overlying shell refuse' (ref. 4). In Fig. 2 the vertical profile revealed by Bishop's excavations is compared with that of the recent investigations. It should be noted also that to the south-west of Cnoc Sligeach the storm-beach debris forms a narrow bar linking two rock knobs. Here the crest of the cobble and boulder debris reaches altitudes up to ~ +11.6 m OD. Collectively the evidence from locations on the south-western, the southern and the south-eastern periphery of Cnoc Sligeach suggests that the site was first occupied about the time of the Holocene marine maximum

or at what was only a short time after the maximum, when severe storms still were building cobble ridges up to levels of ~ +10 to +11 m OD.

The data presented above demonstrate that it is not possible to determine the age relationships between three of the known occupation sites (the Priory Midden and Caisteal nan Gilleann I and II) and the shore line at the (local) maximum of the Holocene marine transgression solely on the basis of stratigraphical and altitudinal evidence. They also suggest that the Cnoc Sligeach site may have been occupied slightly earlier than the Cnoc Coig site. Clearly another basis must be used to determine the age relationships among the five sites themselves and between each individual site and the shore line of the Holocene marine maximum.

Radiocarbon age determinations on occupation material are available from three of the sites (Table 1). The data point to the approximate contemporaneity of the occupation sites at Caisteal nan Gilleann II and Cnoc Coig; statistically no age difference can be established between Birm-347 on the one hand and Q-1353 and Q-1354 on the other (test used: $\Delta A \leq 2\sqrt{(\sigma A_1^2 + \sigma A_2^2)}$). The place of the Cnoc Sligeach site in the occupation chronology is uncertain as yet. This is because one of the three radiocarbon-dated samples was composite in nature (BM-670) and the exact provenances of the other two are not known (GX-1903 and GX-1904). Recent investigations yielded suitable material for radiocarbon dating from accurately identified positions at the Priory Midden, Caisteal nan Gilleann I and Cnoc Sligeach. When the samples have been dated it should be possible to establish the chronology of the sites in relation to each other and the former shore line by combining the evidence of radiocarbon dating with that provided by stratigraphical relationships and altitudinal positions.

I thank Dr P. A. Mellars for many helpful discussions, NERC and the University of Glasgow for meeting field-work expenses and Lord Strathcona for permission to work on Oronsay.

W. G. JARDINE

Department of Geology,
University of Glasgow,
Glasgow, UK

Received 28 February; accepted 18 March 1977.

¹ Mellars, P. & Payne, S. *Nature* 231, 397-398 (1971).

² Jardine, W. G. *Scott. J. Geol.* 9, 214 (1973).

³ Jardine, W. G. in *Quaternary Studies* (eds Suggate, R. P. & Cresswell, M. M. 164 (The Royal Society of New Zealand, Wellington, 1975).

⁴ Bishop, A. H. *Proc. Soc. Antiq. Scotl.* 48, 68 (1914).

Spontaneous formation of bubbles in gas-supersaturated water

SPONTANEOUS nucleation of bubbles may occur in water subjected to heating¹⁻³, tensile stress⁴⁻⁶, or gas-supersaturation⁷⁻¹². Large discrepancies exist between the threshold nucleation conditions predicted theoretically and those indicated by empirical observations¹³. The theoretical approaches have been hampered by inadequate understanding of many of the submicroscopic and molecular properties of liquids, and experimental obstacles have included the always present and difficult to control interference from pre-existing nuclei. These problems have been most pronounced for nucleation induced by dissolved gas; indeed, nucleation thresholds for any gas/liquid system are unknown due to a lack of relevant empirical data, and no quantitative theories seem to exist.

Recent investigations of bubble formation at glass-water interfaces have indicated some minimum nucleation threshold limits for various gases in bulk water^{11,12}. More direct determinations of threshold values, obtained by cinemicrography, are reported here. The results raise some questions about the conventionally accepted relationship between surface tension and pressure in bubbles and the submicroscopic behaviour of water.

In the method used, water equilibrated at high argon gas pressures was decompressed to ambient, atmospheric pressure while it was observed for bubbles. The apparatus (Fig. 1) consisted of a 25 ml chamber for equilibration, A, connected to a 0.5 ml cuvette chamber, B, formed between

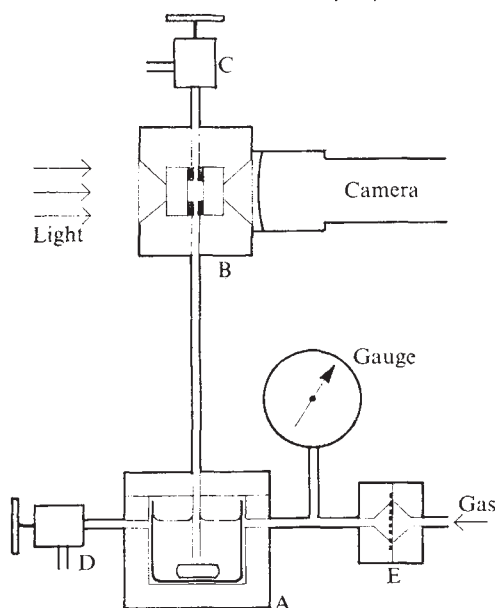


Fig. 1 Schematic diagram of the apparatus which consists of an equilibration chamber A containing a Pyrex cup with a magnetic stirring bar, a cuvette chamber for observations B, valves C and D and Millipore filter unit E. All parts in contact with water were made of stainless steel. A quartz lamp with a light-diffusing, heat-absorbing filter provided illumination.

optical type Pyrex windows (0.64 cm thick, 2.5 cm in diameter) separated by a flat stainless steel ring (0.30 cm thick, 1.40 cm inner diameter). The water in the transilluminated cuvette chamber could be photographed with a 16 mm, 64 frames s⁻¹, movie camera, equipped with an inverted lens of 2.5 cm focal length and an aperture of *f*: 11; the lens to film distance was approximately 18 cm. The full frame coverage was about 0.14 × 0.10 cm, with about 0.02 cm usable depth of field. Before use, all important parts of the apparatus were dichromate-sulphuric acid cleaned and well rinsed. Double glass-distilled water was filtered through a

Millipore filter, pore size 0.45 μm, before equilibration with argon of 99.99% purity. The gas was admitted to the chamber through the same type of filter. Some comparison experiments were performed without filtration.

When gas pressure is applied initially to the system, the water, still with little gas in it, is hydrostatically pressurised; this forces many potential pre-existing gas nuclei into solution. After equilibration at argon pressures ranging from 140 to 202 atm, the water was flushed under pressure through the cuvette chamber by slow and repeated copious bleedings at the valve on the top of the chamber outlet tube, with the assembly in the upright position as shown. The assembly was then turned to the horizontal position, and placed in a predetermined location relative to the camera lens, which had been focused on the nearest inner glass surface. Decompression to ambient pressure was performed over a period of 2–3 s with the camera running. A total of 27 such experiments were performed at room temperature (20–22 °C).

The films were examined frame by frame with ×20 magnification. Bubbles could be detected at 5 μm diameter; those which arose in the water usually could be distinguished easily from those in focus and attached to the glass surface, as the former would usually be out of focus to varying degrees and would show some motion in unison due to a slight inevitable movement of the water caused by the volume expansion of the bubbles.

With the water saturated with argon at 155 atm and lower pressures, bubbles formed only occasionally within the relatively narrow field of the camera, and then only on the glass surface and at decompression pressures close to ambient value. At 158 atm saturation, some bubbles consistently formed on the glass surface. A few also could be observed in the bulk water, but there was some uncertainty as to whether they originated there. With progressively higher saturations the number of bubbles increased, with the largest increase by far occurring in the bulk water, and the bubbles started to form at higher decompression pressures, the higher the saturation. There was a notable synchronous appearance and even distribution of the bubbles in the water, with no coalescence in the early stages. At 165 atm, substantial numbers of bubbles (20–40 per μl) quite clearly developed in the bulk water. This was even more evident during decompression from higher saturations: at 177 atm, the majority of the bubbles (several hundred per μl) appeared in the water and, at 202 atm, at the time bubbles massively occurred in the water (several thousand per μl), relatively few bubbles could be observed on the glass surface.

Because of their presence in such large numbers, their synchronised appearance, and their even distribution, the bubbles observed in the water in general must be nucleated by a spontaneous, homogeneous process involving only the water and the gas dissolved in it, rather than arising from preformed gaseous nuclei in suspended motes or the interface of the motes *per se*. This is supported by the lack of any obvious effect when micropore filtering was eliminated, and also by the lack of any effect on the general cavitation behaviour of water after 1,000 atm hydrostatic prepressurization¹².

In the lower saturation range used here, where the number of bubbles is relatively small, motes may conceivably play a role. Therefore, it is difficult to establish with precision the minimum saturation which is necessary for spontaneous cavitation. It is difficult also to know accurately the supersaturation (that is, saturation pressure minus decompression pressure) which exists in the water at the moment of molecular nucleation, both because of the lack of detailed pressure measurements during decompression and the time delay between the moment of nucleation and the moment of first detection. However, considering all of these factors, the supersaturation for spontaneous nucleation by

argon in water is at least 150 atm, and probably between 155 and 160 atm. These values are the apparent supersaturations directly based on experimental readouts. The true gas tensions in the water may be significantly lower due to the increase in gas solubility which occurs when the pressure is released during decompression. This effect varies with the various gases¹⁴⁻¹⁸ and is not well known for argon. It is, therefore, difficult here to apply the proper correction for the hydrostatic pressure component in the effect.

The nucleation threshold value obtained here is about 30 atm higher than the supersaturation which can be maintained without any cavitation taking place either at glass interface or in the water¹¹. It is within the 100–1000 atm range suggested for spontaneous cavitation in water⁸, but below the value of 250 atm which has been implied from limited experimental observations with oxygen¹⁹.

Although cavitation in the bulk water so far has been directly examined for argon only, the cavitation properties for various gases at the glass–water interface were found to be dependent on the solubility of the gas, and hence its concentration rather than its tension^{11,12}. Thus, the onset of heavy interface cavitation occurred at identical supersaturations (that is, 145–150 atm) for argon and oxygen, two gases with very similar solubilities. For the less soluble nitrogen, 180–190 atm supersaturation was required for the same degree of cavitation. For helium, which has the lowest solubility in water, the value was in excess of 320 atm. Preliminary results on methane and nitrogen obtained by the present technique indicate that the same type of solubility relationship applies directly to the nucleation thresholds in bulk water.

Even a simplified analysis of the present results reveals some highly significant features of the nucleation process. At the critical saturation for spontaneous nucleation, the gas pressure of the nuclei, which would effectively equal the gas tension, must exceed the inward, collapsing force of the surface tension. The critical bubble radius (R) may be expressed (in cm) as: $R = 2\gamma/\Delta P$, where γ is the surface tension (dyn cm⁻¹), and ΔP is the difference (in μ bar) between the hydrostatic and the internal gas pressure. Using this conventionally accepted Laplace formula and a 'normal' surface tension (72.8 dyn cm⁻¹ at 20 °C), the minimum diameter of a stable or expanding nucleus would be 1.8×10^{-6} cm at 160 atm. Such a nucleus would contain at least 10^4 molecules, conservatively estimated by dividing the sphere volume (3.1×10^{-18} cm³) by the cube of intermolecular distance in the compressed gas phase (2.5×10^{-22} cm³). Considering the very dilute concentration of the dissolved gas at this pressure, only about 2×10^{-4} mol dm⁻³, the nearly instantaneous formation of such large gas bodies would be highly improbable, if not impossible. Aggregations brought about by fusions of unstable nuclei conceivably could reduce to some degree the dimensions for critical bubble size, but alone they could not significantly alleviate the basic problem. Also, the spontaneous formation of 'vapour' nuclei of the magnitude indicated above by the self-dispersion of water molecules must be dismissed even though void cavities involving only a few molecules have been postulated to exist in water (refs 9, 20–22).

How then does nucleation occur at the observed supersaturation values? Although it is as yet not possible to propose a quantitative model, or even to describe all of the factors involved, it appears that the surface tension at the gas–water boundary of the nucleus must be at least an order of magnitude lower than that obtained by extrapolation from macroscopic dimensions. The interaction between the compressed gas in the nucleus and the enveloping water would contribute to such a lowering²³, as would the phase boundary curvature effect²⁴, but these probably could not account for all of the effect. A less important factor which conceivably could aid the expansion of some small nuclei (and the collapse of others) would be the 'pres-

sure' fluctuations in their gas phase which consists of relatively few gas molecules. However, due to the ready diffusion of gas, no long-term imbalance could exist between the gas pressure and the gas tension in the water.

This work was supported by the National Institutes of Health, US Department of Health, Education and Welfare.

EDVARD A. HEMMINGSEN

*Physiological Research Laboratory,
Scripps Institution of Oceanography,
University of California, San Diego,
La Jolla, California 92093*

Received 8 January; accepted 9 March 1977.

- ¹ Kenrick, F. B., Gilbert, C. S. & Wismer, K. L. *J. phys. Chem.* **28**, 1297–1307 (1924).
- ² Briggs, L. J. *J. appl. Phys.* **26**, 1001–1003 (1955).
- ³ Reid, R. C. *Am. Sci.* **64**, 146–156 (1976).
- ⁴ Fisher, J. C. *J. appl. Phys.* **19**, 1062–1067 (1948).
- ⁵ Briggs, L. J. *J. appl. Phys.* **21**, 721–722 (1950).
- ⁶ Apfel, R. E. *J. acous. Soc. Am.* **49**, 145–155 (1971).
- ⁷ Kenrick, F. B., Wismer, K. L. & Wyatt, K. S. *J. phys. Chem.* **28**, 1308–1315 (1924).
- ⁸ Harvey, E. N., Barnes, D. K., McElroy, W. D., Whiteley, A. H., Pease, D. C. & Cooper, K. W. *J. cell. comp. Physiol.* **24**, 1–22 (1944).
- ⁹ Fürth, R. *Proc. Camb. Phil. Soc.* **37**, 252–280 (1941).
- ¹⁰ Ward, C. A., Balakrishnan, A. & Hooper, F. C. *J. Basic Engng* **92**, 695–704 (1970).
- ¹¹ Hemmingsen, E. A. *J. appl. Phys.* **46**, 213–218 (1975).
- ¹² Gerth, W. A. & Hemmingsen, E. A. *Z. Naturforsch.* **31a**, 1711–1716 (1976).
- ¹³ Skripov, V. P. *Metastable Liquids* (John Wiley & Sons, New York, 1974).
- ¹⁴ Wiebe, R., Gaddy, V. L. & Heins, C. Jr *J. Am. chem. Soc.* **55**, 547–553 (1933).
- ¹⁵ Wiebe, R. & Gaddy, V. L. *J. Am. chem. Soc.* **57**, 847–851 (1935).
- ¹⁶ Enns, T., Scholander, P. F. & Bradstreet, E. D. *J. phys. Chem.* **69**, 389–391 (1965).
- ¹⁷ O'Sullivan, T. D. & Smith, N. O. *J. phys. Chem.* **74**, 1460–1466 (1970).
- ¹⁸ Schröder, W. Z. *Naturforsch.* **24b**, 500–508 (1969).
- ¹⁹ Clare, N. D. *Proc. R. Soc. Can.* **19**, 32–33 (1925).
- ²⁰ Döring, W. Z. *Phys. Chem. (Leipzig)* **38**, 371–386 (1937).
- ²¹ Finch, R. D. *Phys. Fluids* **12**, 1775–1778 (1969).
- ²² Apfel, R. E. *Scient. Am.* **227**, 58–71 (1972).
- ²³ Eriksson, J. C. *Acta chem. Scand.* **16**, 2199–2211 (1962).
- ²⁴ Buff, F. P. *J. chem. Phys.* **23**, 419–427 (1955).

X-ray analysis of Co–C bond cleavage in the crystalline state

DURING serial structure analysis of cobaloxime derivatives¹, aimed at interpreting the catalytic capability of the asymmetric hydrogenation²⁻⁵, we have found that the crystal of *R*- α -cyanoethyl (*S*- α -methylbenzylamine)-cobaloxime changes its unit cell dimensions by X-ray exposure without degradation of a single crystal form. The rate of the change was so slow that it was possible to collect the intensity data for several intermediate stages. We have proved, by calculating electron density in each stage, that the change reflects the racemisation of the cyanoethyl group.

Figure 1 shows the changes of unit cell dimensions a , b , c and β with exposure time to the conventional MoK α radiation. The cell dimensions at 2 h after the initiation of the irradiation are $a = 8.665(2)$, $b = 13.485(3)$, $c = 9.584(3)$ Å, $\beta = 96.95(3)^\circ$ and $V = 1111.6(5)$ Å³. The a and c axes contract slightly whereas the values of b and β increase to a considerable extent. The rate of change is rapid at very initial stage, but becomes moderately slow. They converge gradually to 8.637(2), 13.833(3), 9.539(2) Å, 99.07(3)° and 1125.5(4) Å³, respectively. After 22 d the changes are within the experimental errors. The space group, P2₁, remains unaltered. The reproducibility was confirmed by runs using three different crystals. The same results were obtained for the crystal of *S*- α -cyanoethyl (*R*- α -methylbenzylamine)-cobaloxime, which is the enantiomer of the present complex.

During these experiments, the appearance of the crystals gave no evidence of decomposition or degradation and the width of each reflection did not change appreciably. The peak and integrated intensities, however, varied to a considerable extent; they fell to an average 50% and 85% of initial values respectively in intermediate stages but fully recovered at the final stage. If the irradiation was interrupted, the change of the cell dimensions became undetectable. By resuming the irradiation, it continued as before.

The three-dimensional intensity data were collected at the stages denoted as A, B, C, D, E, F in Fig. 1. Before A, intensity data collection was unsuccessful. Figure 2 shows the crystal structure, projected along the c axis, deduced from the data at stage A. Fairly loose contacts are observed around the methyl of cyanoethyl group.

The electron density for stage B indicated that the occupancy of the methyl of the cyanoethyl group decreased significantly and a new peak appeared. Figure 3a shows a part of the composite difference Fourier map. The new peak is about twice as high as that of the hydrogen atom and is comparable to the depth of the trough at the position of the methyl group. Between stages B to F, the peak grows higher, and the trough becomes deeper. Figure 3b shows part of the composite electron density map at the final stage

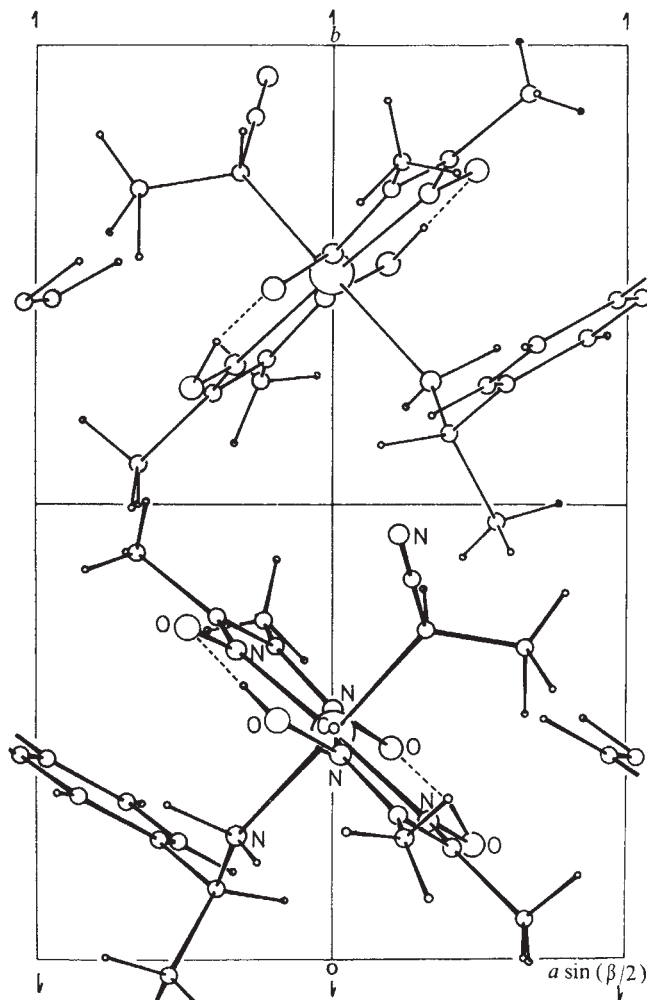


Fig. 2 The crystal structure at the stage A, projected along the c axis. Intensity data for independent reflections were collected with 2θ - ω scan, a scan rate of $8^\circ (2\theta)\text{min}^{-1}$ and background counts of 5 s on both sides of the peak. Because 30 h were required for data collection, the orientation matrix of the crystal was redetermined for several times. A total of 2,069 reflections were measured within the limit $2\theta = 50^\circ$, of which 129 reflections for $|Fo| < 3\sigma(|Fo|)$ were considered unobserved. The structure was solved by the heavy atom method and refined by the block-diagonal least-squares. Final R was 0.055 for 1,940 reflections. The structures for other stages were determined as described above.

F. The height of the new peak is approximately the same as that of the methyl group.

These observations indicate that some portion of the methyl groups is transferred to the position of the new peak so that disordering occurs with the methyl group. The position of the new peak corresponds approximately to that of the methyl group if the absolute configuration of the cyanoethyl group is converted from R to S with the cyano group fixed, although distorted from the ideal geometry because of the repulsion with the equatorial ligands. The least-squares calculations showed that the occupancy factors of R and S methyl groups are 0.75 and 0.25 respectively at the stage B and became equal at the final stage F.

In addition to the disordering of the methyl group, the molecule undergoes gradual translation. At the final stage it moves apart from the original position by 0.04 Å in the a and c directions. No significant changes were observed in bond distances and angles between the structures at the stages A and F.

Although the mechanism of the present transition in the crystalline state is not yet established, it is clear that any of

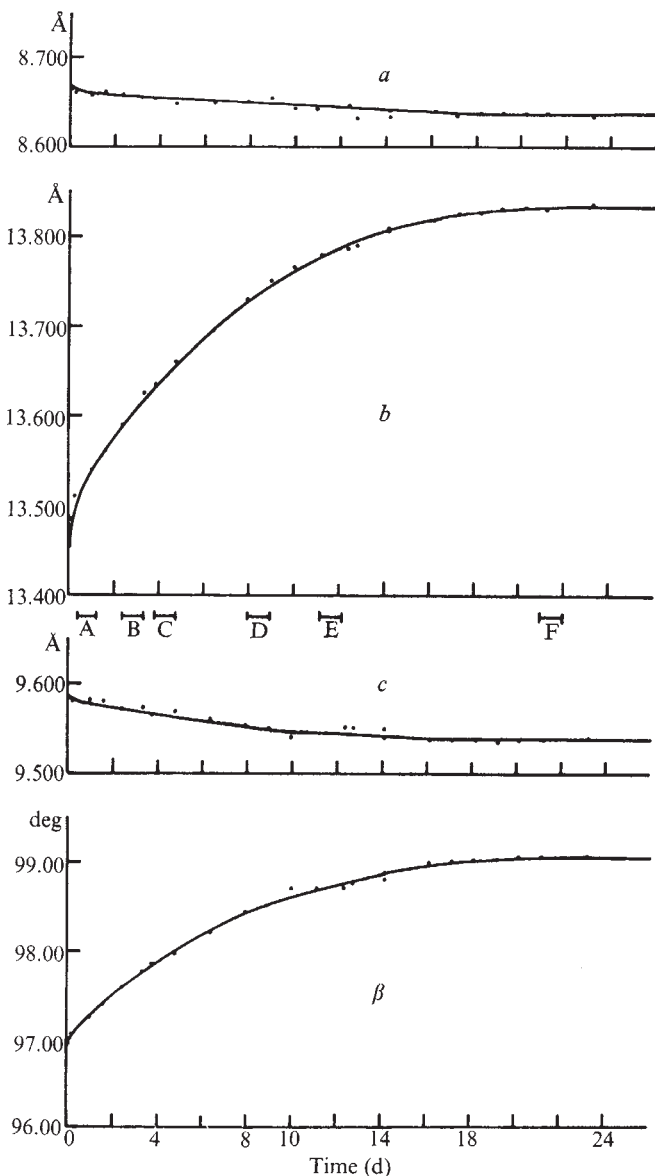


Fig. 1 The change of the unit cell dimensions. A dark red crystal, $0.2 \times 0.2 \times 0.4$ mm, was mounted on a Rigaku automated four-circle diffractometer using $\text{MoK}\alpha$ radiation (40 kV, 20 mA; $\lambda = 0.71069$ Å) monochromatized by graphite. Cell dimensions were determined by least-squares analysis with 15 high-angle ($2\theta > 20^\circ$) reflections. Each value is shown with a point. The experiments were done at intervals of 2 h in the initial stage and done every day in the intermediate stages. By exposure over 22 d, each value fluctuated within its experimental error. E.s.d.s are 0.002–0.004 Å and 0.03–0.05°. At the initial stage the density calculated for $Z = 2$ is 1.388 g cm^{-3} , whereas the density observed by flotation is 1.388 g cm^{-3} . Intensity data were collected at six different stages denoted as A, B, C, D, E, F.

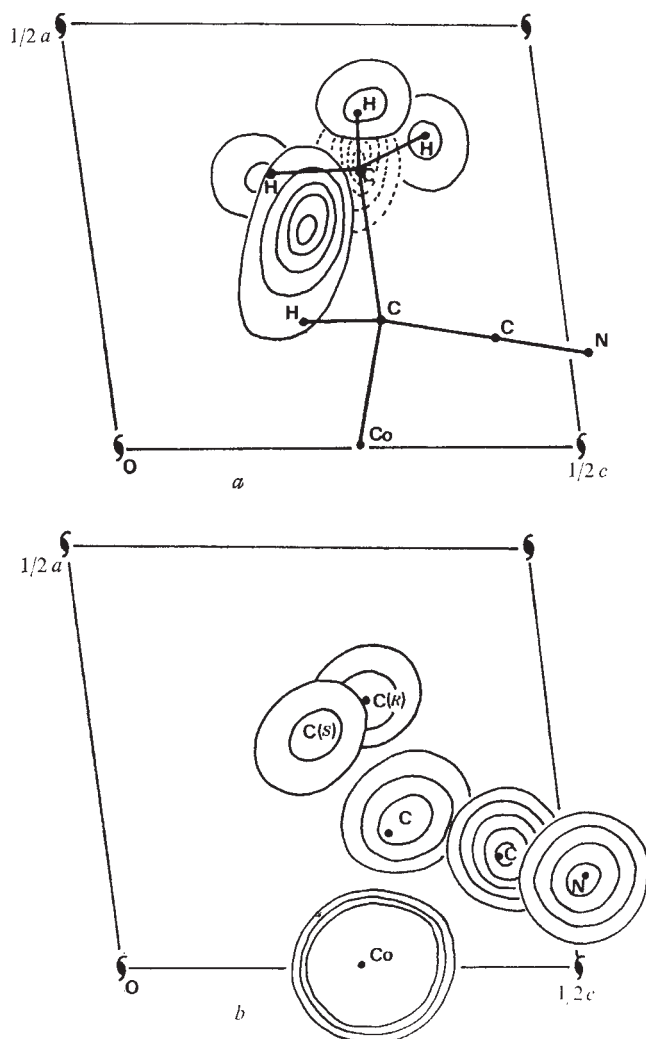


Fig. 3 *a*, Composite difference Fourier map along the *b* axis. Contour interval $0.1 \text{ e} \text{ \AA}^{-3}$; negative contours broken. The analogous map was obtained at each of other stages, though the positive and negative peaks of the new peak and the methyl group respectively are enhanced and the hydrogen peaks are more obscure than those in the figure. In F_c calculation, the carbon atom of the methyl group is taken to have the occupancy factor of 1.0 and the hydrogen atoms of the methyl group are not included. The structures at the stage A are shown with the solid lines for comparison. *b*, The composite electron density map. Contour interval $1 \text{ e} \text{ \AA}^{-3}$; inner contours of the cobalt atom are omitted for clarity. The black circles show the position of the corresponding atoms at the stage A. The asymmetric carbon and nitrogen atoms seem, from the spread and height of the peaks, to be slightly disordered.

four bonds, Co–C, NC–C, H_3C –C and H–C, would be cleaved in the transition; otherwise the asymmetric carbon of *R* configuration could not be converted to *S*. Schrauzer *et al.*⁶ and Giannotti *et al.*^{7–10} showed from the e.s.r. spectra that the irradiation of visible light ($\lambda > 420 \text{ nm}$) results in the homolytic cleavage of Co–C bond in the alkylamine-cobaloxime. It seems quite plausible that for the present complex the same reaction should occur in the secondary process initiated by X rays. Energy released by the bond cleavage might be given to the crystal lattice and contribute to causing topical melting. The cyanoethyl radical would be permitted to rotate to face the opposite side of the plane to the cobalt. This would result in the inversion of the configuration of the cyanoethyl group when the radical makes a bond to cobalt again. As it is quite probable that the racemic structure is more stable thermodynamically,

the growth of the racemic structure would continue without destroying a single crystal form—the racemers having almost identical cell dimensions—so that a whole crystal is converted into the racemic structure.

YUJI OHASHI

YOSHIO SASADA

*Laboratory of Chemistry for Natural Products,
Tokyo Institute of Technology,
O-okayama, Meguro-ku,
Tokyo 152, Japan*

Received 5 January; accepted 4 March 1977.

- 1 Ohashi, Y. *et al. Bull. Chem. Soc. Jap.* **46**, 2589–2590 (1973).
- 2 Ohgo, Y., Takeuchi, S. & Yoshimura, J. *Bull. Chem. Soc. Jap.* **44**, 283–285 (1971); **44**, 583 (1971).
- 3 Takeuchi, S., Ohgo, Y. & Yoshimura, J. *Chem. Lett.* 265–266 (1973).
- 4 Ohgo, Y., Takeuchi, S., Natori, Y. & Yoshimura, J. *Chem. Lett.* 33–36 (1974).
- 5 Ohgo, Y., Natori, Y., Takeuchi, S. & Yoshimura, J. *Chem. Lett.* 709–712 (1974); 1327–1330 (1974).
- 6 Schrauzer, G. N., Lee, L. P. & Sibert, J. W. *J. Am. chem. Soc.* **92**, 2997–3005 (1970).
- 7 Giannotti, C. & Bolton, J. R. *J. organometal. Chem.* **80**, 379–383 (1974).
- 8 Giannotti, C., Merle, G., Fontaine, c. & Bolton, J. R. *J. organometal. Chem.* **91**, 357–362 (1975).
- 9 Giannotti, C., Merle, G. & Bolton, J. R. *J. organometal. Chem.* **99**, 145–156 (1975).
- 10 Giannotti, C. & Bolton, J. R. *J. organometal. Chem.* **110**, 383–388 (1976).

A physicochemical mechanism for magnetic field detection by migratory birds and homing pigeons

MIGRATORY birds and homing pigeons can apparently obtain directional information from the Earth's magnetic field. The effect is difficult to detect, and discussion of the possible process of magnetic field detection by birds seems so far to have foundered on the simple fact that the orientational effect of the Earth's magnetic field on a single electron spin associated with a molecule of animal tissue would be of the order 10^{-8} eV —almost certainly too small to be detectable biologically. Here I direct attention to a process which would overcome this basic problem, and which also seems to provide an explanation of all the main features of published data. It is a mechanism in principle only, however, and is discussed here in no more detail than is necessary to clarify the basic ideas and to provide a basis for further investigation.

There are three requirements that any proposed physicochemical mechanism for field detection must satisfy. First the mechanism must overcome the minuteness of the orientational energy produced by a magnetic field as small as that of the Earth's. Second, the mechanism must be basically unlikely, to account for the weakness of the observed response. Third, if the response is to display the necessary anisotropy (so that the bird prefers one direction to another), detection at the molecular level must take place in an ordered array of molecules, each with an anisotropic response to an applied magnetic field.

A mechanism which satisfies these requirements represents an application of optical pumping which depends on the presence and detection of radiation in the visible region of the spectrum. I propose therefore that magnetic field detection in birds takes place in the eye, in the molecules of the retina (a comparatively well-ordered structure spatially) as an adjunct to the normal processes of vision. Specifically, I propose an optical/r.f. (radio frequency) double resonance process involving the lowest excited molecular triplet state of (say) the rhodopsin molecule. The normal visual process begins with the absorption of a photon in the visible region. For various molecules such

photon absorption is a singlet-singlet transition between the molecular ground state and a vibronically broadened singlet excited state³. Neither of these states, being singlet ($S = 0$), have a magnetic moment, but the triplet states do, having total spin $S = 1$. For such states the anisotropy in the spin-spin dipolar interaction lifts the spin degeneracy in a manner most conveniently characterised by the effective spin-Hamiltonian parameters D and E (refs 4, 5). In conditions of continuous irradiation these triplet levels are dynamically populated through inter-system crossing from the singlet levels through spin-orbit interaction⁶, and experiment has shown that the relative populations of the three triplet states may be appreciably different from unity, depending on the details of symmetry and the relative probability of radiative transitions to the ground state^{6,7}. The relevance of this is that (1) the energy variation of triplet states with field is anisotropic, depending on field magnitude as well as the field direction relative to the principal axes of D and E ; (2) radiative transitions from these states back to the ground state are forbidden transitions and therefore weak; and (3) such transitions have definite polarisation relative to the molecular axes, and the intensity of each transition would depend on the relative populations of the triplet states. If, therefore, the next stage of the detection process were able to monitor radiative intensity at a particular polarisation, and if the relative populations of the triplet states were to vary with magnetic field direction, the detection process would be possible in principle. For this latter requirement we have to postulate first that the parameters D and E produce the necessary magnetic anisotropy, and second, that at least one of the anisotropic $S = 1$ splittings comes into resonance with some internal relaxation process having a frequency around 3 MHz (to satisfy $h\nu_{\text{rf}} = g\beta B$, with $B \approx 10^{-4}$ T). It would be premature to try to specify the possibilities for this at present; optical pumping is a double resonance process and therefore depends not only on the existence of such a resonance but also on various other relaxation rates happening to be within a suitable range. Typical values for D and E for various molecules already studied¹⁻⁶ are two to three orders of magnitude too large, though the requirement here is merely that E should be negligible compared with splittings that would be produced by the Earth's field, while D could take any value. Radiationless decay rates vary widely, but tend to be too fast rather than too slow. The model does therefore face some considerable difficulties in detail, though direct evidence for rhodopsin, or whatever molecule is active in the living creature, is lacking at present. One might therefore have been reluctant to postulate that birds were able to detect the Earth's field, but here one is faced with the fact that they do. I now turn to a review of the experimental data in the light of the above model.

The main findings seem to be: (1) response to field is evident over a wide range of illumination—daytime overcast¹¹ to starlit sky¹⁷—but not, so far as I am aware, in total darkness (see (7) below). This is consistent with a detection process based on the normal visual process. (2) The response is much weaker than other directional clues (though stronger, apparently, than stellar constellation information¹⁷) and is difficult to detect experimentally. This is consistent with the electric-dipole forbidden nature of triplet state radiative decay, and the probability of competing radiationless decay processes. (3) The response to magnetic field is axial rather than polar. That is, birds show sensitivity to the angle of dip rather than the N-S direction of the field⁸⁻¹⁰. This is consistent with a resonance process (though there are slight difficulties concerned with the detection of a particular polarisation in the next stage of the process) and is perhaps the best evidence there is for the proposed model, particularly when taken together with the fact that (4) the response to magnetic field reaches a peak for fields comparable with the Earth's, falling off at both lower and higher fields⁸. (5) The response to field disappears temporarily if the field is altered a little ($\sim 20\%$), but returns if the birds remain in the altered field for a day or two⁹. This is consistent with the detection process taking place in a directionally ordered

structure (such as the retina) and where the molecular response to field is axial; that is only the field component parallel to the molecular axis matters, as in the case $D \gg E$. In this case for a particular eye orientation the resonance condition would be satisfied on particular areas of the retina. Altering the field would shift the resonance condition to some other region of the retina and confuse the bird initially. A day later, however, it could have learnt to respond to the altered cue. (6) Birds and a large number of other creatures are able to detect polarisation of light¹². This need not necessarily have its origin in triplet state radiative decay, but it is a requirement of the proposed model. (7) Attempts to operantly condition pigeons to magnetic field have so far failed to produce unambiguous results^{13,14}. A possible reason for this may be found (thanks to Professor Keeton for providing further details of experiments from ref. 14) in the lighting level during the experiments. Birds go to sleep in total darkness⁹, while normal laboratory lighting is too bright for operant conditioning, so the light level is therefore used to set the bird's level of awareness to a hopefully suitable level for the experiment. If, however, field detection does take place in the eye, there is some possibility of 'tuning out' precisely the effect one is seeking, by using the light level in this manner. (8) There is somewhat sketchy evidence that birds lose their orientation in the vicinity of radar installations¹⁵. This is consistent with resonances in the GHz region due to the D splitting of the triplet state, which would tend to disturb the equilibrium due to the internal r.f. resonance as postulated.

Thus, in conclusion, optical pumping overcomes the basic problem faced by any proposed model for field detection biological systems, and also the excited triplet state provides for the anisotropy which is essential if such detection is to be useful directionally. Though the requirements placed on the parameters D and E and all the appropriate relaxation processes are rather stringent, the straightforward explanation which this model provides for the existing data does provide some grounds for optimism. The questions it raises are obvious—is vision essential to magnetic field detection? Can the physico-chemical requirements be satisfied? Behaviourally, it might be interesting to conduct the directed-perching type of experiment⁹ in monochromatic light at long wavelengths, in the hope of modifying the excitation transfer probability to the lowest excited triplet state. Finally, nothing in the proposed mechanism is specific to birds, rather than other creatures which have likewise shown sensitivity to the Earth's magnetic field¹⁶.

I thank Professors G. V. T. Matthews, Sir Denys Wilkinson and W. T. Keeton for discussions and encouragement. Dr D. T. Edmonds provided useful ideas on relaxation processes, and Miss M. M. Davidson assisted in the literature research. I am also grateful to the British Broadcasting Corporation, whose Horizon programme on Professor Keeton's researches first stimulated my interest in the problem.

M. J. M. LEASK

Clarendon Laboratory,
Parks Road, Oxford, UK

Received 23 November 1976; accepted 16 March 1977.

- ¹ Matthews, G. V. T. *Orientation and Position Finding by Birds* (Oxford University Press, Oxford, 1973).
- ² Matthews, G. V. T. *J. R. Stat. Soc. B* **36**, 349-364 (1974).
- ³ Wile Robinson, G. *Excited States* **1**, 1-34 (1974).
- ⁴ Ayscough, P. B. *Electron Spin Resonance in Chemistry*, ch. 11 (Methuen, London, 1967).
- ⁵ Abragam, A. & Bleaney, B. *Electron Paramagnetic Resonance of Transition Metal Ions*, ch. I (Oxford University Press, Oxford, 1970).
- ⁶ El-Sayed, M. A. *Excited States* **1**, 35-77 (1974).
- ⁷ Moore, T. A. & Kwiram, A. L. *Biochemistry* **13**, 5403-5407 (1974).
- ⁸ Wiltchko, W. & Wiltchko, R. *Science* **176**, 62-64 (1972).
- ⁹ Wiltchko, W. *Animal Orientation and Navigation* (eds Galler, S. R. et al.) 569-578 (NASA, Washington D.C., 1972).
- ¹⁰ Walcott, C. & Green, R. P. *Science* **184**, 180-182 (1974).
- ¹¹ Keeton, W. T. *Animal Orientation and Navigation* (eds Galler, S. R. et al.) 579-594 (NASA, Washington D.C., 1972).
- ¹² Kreithen, M. L. & Keeton, W. T. *J. cell. comp. Physiol.* **91**, 83-92 (1974).
- ¹³ Reille, A. *J. Physiol., Paris* **60**, 85-92 (1968).
- ¹⁴ Kreithen, M. L. & Keeton, W. T. *J. cell. comp. Physiol.* **91**, 355-362 (1974).
- ¹⁵ Yeagley, H. L. *J. appl. Phys.* **22**, 746-760 (1951).
- ¹⁶ Wehner, R. & Labhart, T. *Experientia* **26**, 967-968 (1970).
- ¹⁷ Wiltchko, W. & Wiltchko, R. *Z. Tierpsychol.* **37**, 337-355; **39**, 265-282 (1975).

Sources of sulphur in rain collected below a wheat canopy

VEGETATION plays an important role in the cycle of sulphur between the atmosphere and the soil¹⁻⁴. We have measured the quantity of sulphur in rain collected below a maturing wheat canopy. This sulphur has three sources: first, the atmosphere, from which falling rain gains SO₂ and sulphate; second, leaf surfaces, from which rain washes sulphur which was previously deposited by turbulent transfer ('dry deposition'), and third, leaf tissue, from which rain leaches sulphur⁵. We have now deduced from field and laboratory measurements that leaching supplied nearly 90% of the sulphur gained by rain as it fell through the wheat canopy. Only a small fraction of sulphur which had been dry-deposited on the surface of leaves could be washed off.

The measurements were made in June, July and August 1975 in a 6.9-ha field of winter wheat (variety Maris Huntsman) growing at Sutton Bonington (52.8°N, 1.25°W), a rural area within an industrial region. Rainfall above the canopy was collected at 12 sites in the field in polyethylene funnels, diameter 14 cm, which drained into glass bottles. At each site, throughfall below the canopy was collected in polyvinyl chloride guttering, 2 m × 7 cm, laid along the rows and draining into glass bottles. These linear gauges were calibrated over short grass by comparison with a standard rain gauge. The funnels and gutters were wiped with tissues soaked in distilled water at least every 4 d and whenever possible immediately before rain fell. This practice may lead to underestimation of the quantity of rain received in light showers but it minimises analytical errors from contamination⁶. A linear regression of throughfall, T (mm), against rainfall, R (mm), for 11 events yielded $T = 0.50R - 0.036$ ($r = 0.99$). Experiments with artificial rain indicated that the canopy storage capacity immediately after rain was 0.5 mm. We were unable to measure stemflow, S , so we assumed that it accounted for the residual, that is $R - 0.5 = S + T$. The large stemflow component is consistent with the upright structure of the crop and the compositions of stemflow and throughfall were assumed to be identical because both passed over similar leaf surfaces.

The wheat crop was sampled at weekly intervals by collecting tops from four randomly selected rows, 2 m long, and roots from four soil cores to 2-m depth. The four top samples were digested individually; roots were washed free of soil and digested in bulk⁷. The sulphur contents of digested plant samples and of rainfall were determined by precipitation of barium sulphate. Washing followed by redissolution of the precipitate in EDTA facilitated the measurement of barium by atomic absorption spectrophotometry.

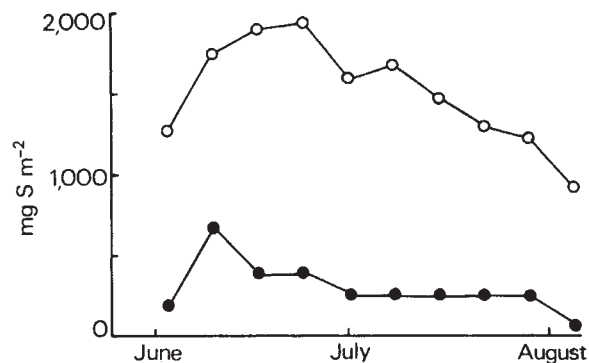


Fig. 1 Sulphur content of winter wheat in 1975 expressed as sulphur per unit ground area ●, Roots; ○, roots plus tops. Anthesis was on 15 June and harvest on 5 August.

Table 1 shows that after rain additional sulphur was always found in water collected below the canopy. The excess consists of dry deposition washed from leaf surfaces and leachate from leaf tissue. The contributions from each source during the whole period can be assessed as follows.

Fowler^{2,8} made a detailed study of dry deposition of SO₂ on wheat growing in the same field in 1974 and deduced typical values of the deposition velocity, v_d , to wet and dry leaf surfaces⁸. The mass, M , of sulphur dioxide deposited in time, t , on leaves covering unit ground area is $M = v_d \chi t$ where χ is the SO₂ concentration above the surface. While our measurements were being made, the mean value of χ at a nearby site was $32 \mu\text{g m}^{-3}$. Meteorological observations indicated that leaves were wet (mainly with dew) for about 15% of the 66-d period. Using Fowler's values $v_d = 1.5 \text{ mm s}^{-1}$ (dry) and 10 mm s^{-1} (wet) we estimate that M corresponded to about 250 mg S m^{-2} during the whole period, that is deposition on the cuticle was similar in magnitude to the 236 mg m^{-2} additional sulphur found below the canopy (Table 1).

To estimate the fraction of deposited sulphur which could later be washed off the cuticle, we exposed leaves of senescent wheat plants growing in pots to radioactive ³⁵SO₂, removed the leaves at intervals ranging from 10 min to 8 d, and washed leaf surfaces thoroughly with distilled water. Stomata in senescent leaves present a large resistance to gas exchange⁹ so we assume that most of the radioactive sulphur was on the cuticle. The fraction of total activity which could be washed off had a mean of 0.13 and ranged from 0.04 to 0.19. If the same mean figure is valid in the field, dry deposition washed from leaves supplied only about 30 mg S m^{-2} of the additional 236 mg S m^{-2} recorded below the canopy (Table 1). Garland and Branson found a similar fractional washoff from pine shoots exposed to ³⁵SO₂ (ref. 10).

Table 1 Sulphur contents of rainfall and throughfall below a wheat canopy

Period of rainfall	Rainfall (mm)	S content of rain ($\mu\text{g ml}^{-1}$)	S input to canopy in rain (mg m^{-2} ground area)	S content of throughfall ($\mu\text{g ml}^{-1}$)	S output from canopy by throughfall and stemflow (mg m^{-2} ground area)
2-3 June	4.3	$3.1 \pm 0.2^*$	13.3	3.6 ± 0.2	13.3
15-16	10.7	3.3 ± 0.3	35.3	5.0 ± 0.4	51.0
8-10 July	1.6	13.9 ± 1.1	22.4	34.3 ± 2.6	37.7
10-13	5.6	3.4 ± 0.2	19.0	24.8 ± 2.6	126.5
13-15	9.1	2.0 ± 0.1	18.2	6.8 ± 0.7	58.5
15-18	16.9	2.2 ± 0.1	37.2	3.5 ± 0.2	57.4
18-21	11.3	3.3 ± 0.1	37.3	6.0 ± 0.4	64.8
21-23	2.9	3.0 ± 0.2	8.7	5.3 ± 0.5	12.7
4-5 Aug	1.3	11.1 ± 0.4	14.4	25.3 ± 2.2	20.2
			Total 205.8		Total 442.1
					Additional sulphur 236 mg m^{-2} ground area

*s.e.m.

Figure 1 shows the results of the sulphur analysis of whole plants and roots expressed as sulphur per unit ground area. Clearly sulphur losses from the plants after anthesis (15 June) greatly exceeded sulphur additions measured in throughfall. The sulphur content of the tops decreased by about 40% after this time, but Table 1 shows that leaching from leaves accounted for only about one third of the sulphur lost. Because sulphur did not accumulate in roots we conclude that it was removed by other mechanisms as well as by leaching in rain. Exudation of sulphur from roots of plants growing in solution and exposed to SO_2 has been reported¹¹⁻¹³. Knowles and Watkin¹⁴ observed large losses of potassium (approximately 50% of maximum content) from wheat plants after anthesis and they postulated an efflux through the roots to the soil. A similar efflux of sulphur seems the most likely explanation of our results. Emission of gaseous sulphur compounds from leaves (reported for tomato plants¹⁵ and pine needles¹⁶) is another possibility.

Our measurements and calculations show that leaching of plant sulphur by rain was the main source of additional sulphur in throughfall, and that most sulphur dry-deposited on leaf surfaces remained fixed there. Current studies in Scandinavia aim to deduce deposition velocities for SO_2 to forests by ascribing part of the excess sulphur in throughfall to deposited sulphur washed from leaf surfaces³. Our results show that this technique would be highly inaccurate for a maturing wheat crop, and it is clearly desirable to make further studies with vegetation of other ages and species to establish the pathways in the sulphur cycle more clearly.

We thank D. V. Crawford, J. A. Garland, J. R. Branson and A. Martin for advice and facilities. The Central Electricity Generating Board provided financial support for C.R.

CLAIRE C. RAYBOULD
M. H. UNSWORTH
P. J. GREGORY

Department of Physiology and Environmental Studies,
University of Nottingham School of Agriculture,
Sutton Bonington, Loughborough, UK

Received 24 January; accepted 9 March 1977.

- ¹ Garland, J. A., Atkins, D. H. F., Readings, C. J. & Caughey, S. J. *Atmos. Env.* **8**, 75-79 (1974).
- ² Fowler, D. & Unsworth, M. H. *Nature* **249**, 389-390 (1974).
- ³ Doveland, H., Joranger, E. & Semb, A. *Res. Rep. FRG/76*, SNSF-project (Norwegian Institute for Forest Research, Oslo, 1976).
- ⁴ Cowling, D. W. & Lockyer, D. R. *J. exp. Bot.* **27**, 411-417 (1976).
- ⁵ Tukey, H. B., Jr *A. Rev. Pl. Physiol.* **21**, 305-324 (1970).
- ⁶ Mellanby, K. *Nature* **264**, 167 (1976).
- ⁷ Cunningham, R. K. *Chem. Ind.* 2120-2121 (1962).
- ⁸ Fowler, D. thesis, Univ. Nottingham (1976).
- ⁹ Meidner, H. & Mansfield, T. A. *Physiology of Stomata* (McGraw Hill, New York, 1969).
- ¹⁰ Garland, J. A. & Branson, J. R. *Tellus* (in the press).
- ¹¹ Thomas, M. D. & Hill, G. R. *Pl. Physiol.* **12**, 309-383 (1937).
- ¹² Friend, M. *Proc. Soil. Sci. Soc. Am.* **13**, 135-138 (1948).
- ¹³ Garsed, S. G. & Read, D. J. *New Phytol.* **78**, 111-119 (1977).
- ¹⁴ Knowles, F. & Watkin, J. W. *J. agric. Sci., Camb.* **21**, 612-637 (1931).
- ¹⁵ de Cormis, L. in *Air Pollution, Proc. 1st Eur. Congr. Influence of Air Pollution on Plants and Animals*, 75-78 (PUDOC, Wageningen, 1969).
- ¹⁶ Krouse, H. R. *Nature* **265**, 45-46 (1977).

Seed germination in response to diurnal fluctuations of temperature

DIURNAL fluctuations in temperature may initiate or accelerate germination in certain flowering plants¹⁻⁸, and the effectiveness of the stimulus varies according to the amplitude of fluctuation⁷ and the presence or absence of light⁸. Attempts^{3,8} to assess the adaptive significance of the phenomenon, however, have been limited by the scarcity of data for species of contrasted ecology. We report here an investigation of germination responses to fluctuating temperatures, conducted on seeds of herbaceous species collected from native populations near Sheffield. The results suggest that requirements for diurnal fluctuations in

temperature are characteristic of the germination of species from particular types of habitat and provide mechanisms which cause seeds to germinate at times and in places favourable for seedling establishment.

Species were screened for the capacity to respond to fluctuating temperatures in the light in two stages. First, germination was compared in two incubators. One consisted of a growth room with a diurnal fluctuation of 5 °C (20 °C day, 15 °C night) and the other was a thermogradient bar⁹ providing 'constant' temperatures (maximum diurnal fluctuation 1 °C) for 30 batches of seeds distributed at regular intervals across the range 5-40 °C. This procedure revealed a large number of species which germinated to a high percentage in fluctuating temperatures but which remained partially or totally dormant in constant conditions. Such species included several marsh plants, in some of which germination on the thermogradient bar was restricted to exceptionally high (>25 °C) temperatures.

In the second stage of screening, germination of each of these species was examined in a new incubator⁷ which provides diurnal temperature fluctuations over a selected range of amplitudes. The results described here refer to experiments in which batches, each of 30 seeds, experienced diurnal fluctuations over the amplitude range 0-12 °C. Details of the methods are given in the legend of Table 1. Each fluctuation was applied in the form of a depression

Table 1 Species in which germination in the light is promoted by fluctuating temperatures

Species	Fluctuating temperature requirement (°C)
<i>Agropyron repens</i> *	4.0
<i>Atriplex hastata</i>	6.5
<i>Bidens tripartita</i> *	4.0
<i>Cardamine pratensis</i>	7.0
<i>Carex otrubae</i>	1.0
<i>Chenopodium rubrum</i>	7.0
<i>Filipendula ulmaria</i>	1.5
<i>Gnaphalium uliginosum</i> *	5.0
<i>Juncus effusus</i>	1.5
<i>Lycopus europaeus</i> *	8.0
<i>Mentha arvensis</i> †	4.5
<i>Phalaris arundinacea</i>	1.0
<i>Polygonum persicaria</i>	8.0
<i>Ranunculus repens</i> ‡	1.0
<i>Rorippa islandica</i> *	9.0
<i>Rumex obtusifolius</i>	2.0
<i>Rumex sanguineus</i> *	2.5
<i>Sieglingia decumbens</i> *	1.0
<i>Sonchus arvensis</i>	7.5
<i>Tripleurospermum maritimum</i>	1.5
<i>Typha latifolia</i>	1.0
<i>Urtica dioica</i>	5.5

The figures indicate the amplitude of fluctuation required to attain 50% of the maximum germination recorded in the experimental conditions. Germinated seeds were counted and removed daily. Diurnal fluctuations were maintained throughout each experiment, which was terminated when germination ceased. Light intensity during the light period was 1,846 $\mu\text{W cm}^{-2}$. Before the test all seed was stored dry at 5 °C except for the following.

*20 °C dry storage.

†2 °C wet storage.

‡20 °C wet storage.

Further studies have indicated that requirements for fluctuating temperatures in certain of these species may decline gradually during extended laboratory storage.

below a base temperature of 22 °C extending for 6 in every 24 h and coinciding with a dark period. Such downward fluctuations were preferred so that no regime should impinge on the high temperatures known to initiate germination in certain of the species. Thus any responses could be attributed to the fluctuation itself rather than to experience of particular absolute temperatures.

In Table 2, the 292 species screened are classified arbitrarily into five ecologically defined groups, for each of which the proportion of species responding to fluctuating

temperatures in the light has been calculated. Some of the species stimulated by fluctuating temperatures are identified in Table 1, which also includes the amplitude required to achieve 50% of the maximum germination recorded in the experimental conditions. A high incidence of responses to fluctuating temperatures is apparent in species of marshes, bogs, streamsides and other wetland habitats. The significance of this correlation is strengthened by the fact that several ruderal species in which germination was found to depend on fluctuating temperatures, such as *Chenopodium rubrum*, *Polygonum persicaria* and *Polygonum lapathifolium*, are particularly common in areas of arable fields subject to seasonal waterlogging, and frequently colonise the mud adjacent to lakes, ponds, ditches and reservoirs¹⁰.

Further investigations with modified incubators were conducted to examine the response of seeds to diurnal temperature fluctuations in light and in darkness. Specimen results are shown in Fig. 1. In some marsh plants and ruderals darkness increased considerably the amplitude of fluctuation required to initiate germination. For many species which in light germinated rapidly and to a high percentage in constant temperatures, a large proportion of the seeds required large (>4 °C) fluctuations to germinate in darkness. This category included species from grassland, heathland and a wide range of vegetation types.

Table 2 Comparison of frequency of occurrence of requirements for fluctuating temperature for germination in the light

Habitat	No. of species examined	% Stimulated by fluctuating temperatures
Wetland	68	43
Disturbed ground	61	18
Grassland	100	2
Woodland	20	5
Remainder	43	0

Further experiments are necessary to estimate how far these results from local seed collections are representative of each species. Already, however, responses in two ecologically distinct groups are evident. The first (group 1) consists of species (mainly marsh plants) which require fluctuating temperatures for germination in the light. The second (group 2), including plants from a wide range of habitats, consists of species in which the response to fluctuating temperatures occurs only in conditions of continuous darkness or is greater in darkness than in the light. We suggest that in the two groups responses to temperature fluctuations are initiated by different environmental cues.

We consider first group 1. In temperate regions, seedling establishment in many marshland species is restricted to late spring when the water table recedes, causing bare soil to be exposed and its surface horizons to become aerobic. We suggest that fluctuating temperature requirements delay germination until such time in the spring that the declining water content of the soil causes the seed to be no longer insulated against the increasing daily radiation load.

The necessity of bare soil for seedling establishment is not confined to ruderal species. Ecologists¹¹⁻¹⁴ have noted that seedlings of many group 2 species establish successfully only in gaps in the canopy of leaves and litter. From studies of the seed content of soils¹⁵⁻²⁰ and our own measurements it is apparent that buried seeds of many herbs persist in large numbers beneath closed canopies. Such species include all those which in Fig. 1 display a requirement for fluctuating temperatures in darkness. We suggest therefore that germination is initiated in gaps in vegetation by two processes: (1) unearthing of buried seeds and triggering of germination by exposure to light and (2) germination of buried seeds *in situ*, caused

by large diurnal temperature fluctuations resulting from removal of the insulating effect of foliage and litter. The results of a temperature monitoring experiment in a grass sward (Fig. 2) confirmed that at a soil depth of 1 cm, beneath a closed canopy, diurnal fluctuations in temperature exceeded 4 °C only in exceptionally hot dry conditions, whereas, in neighbouring gaps, amplitudes greater than 10 °C were frequent.

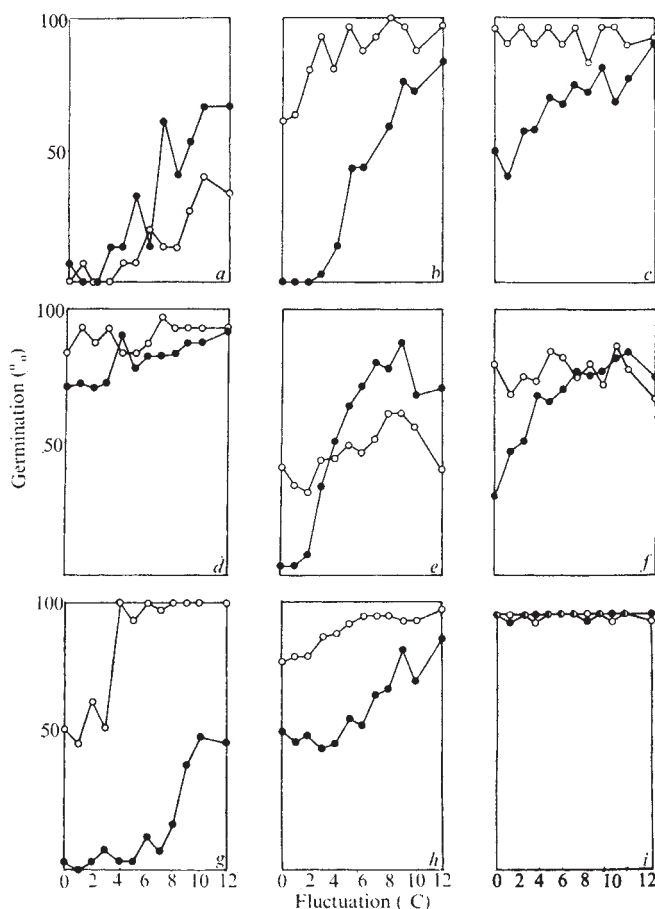


Fig. 1 Germination responses to various amplitudes of diurnal fluctuations of temperature in light (○) and in continuous darkness (●) in nine herbaceous plants. Details of the experimental procedure are provided in the legend of Table 1. The erratic course of the 'curves' for certain species, for example, *A. repens*, probably reflects the heterogeneity of responses within the seed populations studied. In any sample of such seed there exists a mixture of phenotypes differing in response to temperature fluctuations. Such variability could be eliminated only by the use of very large numbers of seeds in the experiments. The results for *A. repens* and *P. pratensis* suggest that even when exposed to large fluctuations of temperature, a proportion of the seeds was inhibited by the light treatment. This suggests a mechanism whereby germination of seeds of these species may be delayed until after burial in the soil. a, *A. repens*; b, *D. caespitosa*; c, *H. lanatus*; d, *P. annua*; e, *P. pratensis*; f, *P. trivialis*; g, *R. obtusifolius*; h, *S. media*; i, *L. perenne*.

Control of arable weeds is complicated by the erratic germination patterns exhibited by certain species. It seems likely that one reason for the low predictability of germination in species such as dock (*Rumex obtusifolius*), couch grass (*Agropyron repens*) and chickweed (*Stellaria media*) is the heterogeneity of their seeds with respect to the amplitude of fluctuation required to induce germination in darkness (Fig. 1). In field conditions such variability is compounded by the fact that seeds are distributed throughout the soil and their exposure to temperature fluctuations will be modified by variations in incident radiation, vegetation canopy and soil moisture content.

Responses of buried seeds to fluctuating temperatures

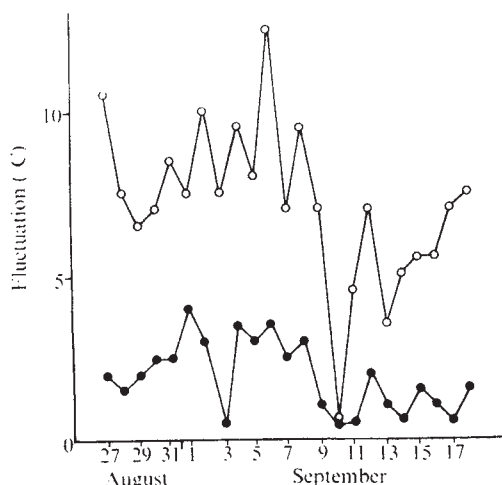


Fig. 2 Comparison of the amplitudes of diurnal fluctuations of soil temperature beneath an established sward (●) and in an artificially created gap of 20-cm diameter (○). Measurements were made at a depth of 1 cm. Recording was carried out during the period 27 August to 18 September 1976. Each point refers to the difference between the minimum and maximum temperatures on the day concerned.

are also relevant to the 'deterioration' of pastures, characterised by a decline of the sown species (usually *Lolium perenne*) and an incursion by native grasses. Figure 1 shows that in native sward-invading species such as *Poa annua*, *P. trivialis*, *P. pratensis*, *Holcus lanatus* and *Deschampsia caespitosa*, a proportion of the seeds require large temperature fluctuations for germination in darkness. It is not surprising therefore that these species occur in the reservoir of buried seeds in pastures¹⁶ and exploit gaps in the canopy arising from damage to the vegetation. The decline of *L. perenne* may be related in part to the fact that it is insensitive to diurnal temperature fluctuations and is not incorporated into the buried seed flora.

We thank Mr A. Curtis, Mr A. Neal, Miss J. Rodman, Miss S. Lee and Mr A. H. Lumb for assistance. This work was supported by the NERC.

K. THOMPSON
J. P. GRIME
G. MASON

Unit of Comparative Plant Ecology,
(Natural Environment Research Council),
Department of Botany,
University of Sheffield
Sheffield, UK

Received 18 January; accepted 17 March 1977.

- Morinaga, T. *Am. J. Bot.* **13**, 141–158 (1926).
- Warrington, K. *J. Ecol.* **24**, 185–204 (1936).
- Cavers, P. B. & Harper, J. L. *J. Ecol.* **54**, 307–382 (1966).
- Thompson, P. A. *J. exp. Bot.* **20**, 1–11 (1968).
- Thompson, P. A. *J. exp. Bot.* **25**, 156–163 (1974).
- Thompson, P. A. *J. exp. Bot.* **26**, 164–175 (1974).
- Grime, J. P. & Thompson, K. *Ann. Bot.* **40**, 795–799 (1976).
- Toole, V. K. & Borthwick, H. A. *J. Am. Soc. Hort. Sci.* **93**(3), 301–304 (1971).
- Haldall, P. & French, C. S. *Pl. Physiol.* **33**, 249–252 (1958).
- Clapham, A. R., Tutin, T. G. & Warburg, E. F., 2nd ed. *Flora of the British Isles* (1962).
- Kershaw, K. A. *J. Ecol.* **46**, 571–592 (1958).
- Milnerthorpe, F. L. *Mechanisms in Biological Competition* 342 (Cambridge University Press, 1961).
- Rorison, I. H. *J. Ecol.* **48**, 585–599 (1960).
- Wardle, P. *J. Ecol.* **47**, 483–497 (1959).
- Chapman, S. S. & Morris, K. *J. Ecol.* **36**, 149–173 (1948).
- Chippendale, H. G. & Milton, W. E. *J. Ecol.* **22**, 508–531 (1934).
- Roberts, H. A. *Rep. natn. Veg. Res. Stn.* 25–38 (1969).
- Sarukhan, J. *J. Ecol.* **62**, 151–177 (1974).
- Strickler, G. S. & Edgerton, P. J. *Ecology* **57**, 801–807 (1976).
- Wesson, G. & Wareing, P. F. *J. exp. Bot.* **20**, 402–413 (1969).

Energy requirement for symbiotic nitrogen fixation

COMPARING the energy required by legumes for symbiotic nitrogen fixation with that of assimilation of nitrate, Gibson¹ concluded the costs to be about the same. About 15% of net photosynthetic production by the plant may be used in meeting its nitrogen requirements. If energy for the fixation of nitrogen symbiotically and that for the assimilation of NH_4^+ or NO_3^- from the soil solution are both provided by the chemical products of photosynthesis, then the CO_2 respired in supplying that energy must contribute to the total efflux of CO_2 from the plant in the dark. McCree² and Thornley³ have discussed methods of partitioning the dark CO_2 efflux into a growth (synthesis) and a maintenance component. Both nitrogen fixation and nitrogen assimilation can be expected to contribute to the CO_2 flux associated with synthesis. If Gibson is correct, nodulated plants using only symbiotically fixed nitrogen should have the same growth coefficient as non-nodulated plants supplied with exogenous mineral nitrogen when grown in the same conditions. I have examined this hypothesis using a modification of the method used by McCree with *Trifolium subterraneum* L. cultivar Woogenellup as test material.

Plants were grown at a density of 2,000 m^{-2} in square, 3-l plastic pots. The rooting medium was a fritted clay of excellent water holding capacity, low bulk density (450 kg m^{-3}) and with no measurable capacity to absorb or evolve CO_2 . Plants in one group of pots (A) were inoculated with *Rhizobium* and watered daily with a nutrient solution free of mineral nitrogen. The other group (B) was not inoculated and was watered with the same nutrient solution but with 7.5 mM NO_3^- partially replacing SO_4^{2-} . Pots were housed in a growth cabinet set at $20 \pm 0.5^\circ\text{C}$ and plants were exposed to a photosynthetic photon flux density of 500 $\mu\text{E s}^{-1} \text{m}^{-2}$ for 12 h each day. Plants subjected to the two treatments grew equally vigorously. After a full canopy had developed, the CO_2 exchange rates were measured in an assimilation chamber (open system) at 20°C and the same light flux density as the growth cabinet. CO_2 flux into and out of the whole pot was measured continuously and the total daytime influx (D) and night efflux (N) were obtained by integration. This was done for three successive 24 h periods, each pot always reaching a steady state of near constant D and N over the second and third periods. At the beginning of the fourth day the light was turned off, the plants remaining in darkness for a further 36 h. The rate of CO_2 efflux declined rapidly over the first 4 h of extended night, reaching a new, steady, low rate at the beginning of the next day: a new total 12-h efflux (N_a) was calculated from this reduced value. N_a was always about half N.

The method of estimating the growth coefficient k in the equation

$$N = kD + (1+k)cW \quad (1)$$

will be described more fully elsewhere. It is assumed^{2,4} that N_a represents the maintenance flux and is equal to cW where W is the total plant dry weight in CO_2 equivalents, and c the maintenance coefficient (mg CO_2 per g CO_2 per 12 h). The growth coefficient k represents the fraction of the net CO_2 uptake in the light (D) that is associated with the synthesis of new dry matter (g CO_2 per g CO_2) over a 12-h period.

$$k = (N - N_a) / (D + N_a) \quad (2)$$

Values of k and c averaged over six replicates were:

	k	c
A (nodulated)	0.189 ± 0.017	11.9 ± 2.0
B (non-nodulated)	0.137 ± 0.010	13.2 ± 1.8

The two values of k are highly significantly different in a t test: those of c do not differ. A value of $k = 0.144$ has been obtained for non-nodulated plants grown elsewhere in the same conditions.

Assuming the dark CO₂ efflux (that is, *N*) to be the same during the 12-h day as during the 12-h night, these results show that plants fixing nitrogen symbiotically from the air used 37.8% (that is 2 *k*) of the net daily CO₂ uptake for synthesis of new material in 24 h. Those assimilating mineral nitrogen from the soil solution used 27.4% in 24 h. Nodulated plants averaged 3.9% nitrogen, the non-nodulated 4.5%, these values being each a weighted mean for the whole plant. Nodulated plants used 810 mg of CO₂ for the synthesis of 1 g of dry weight of plant material whereas for non-nodulated the figure was 510 mg. The latter is slightly higher than the value of 494 mg calculated from data of Penning de Vries⁵ for a proximate composition of 1 g of clover biomass as protein, 0.24; carbohydrate 0.61; lipid 0.07; minerals 0.08. Calculation of the conversion coefficient (*Y_G*)³ by the equation:

$$Y_G = (D - N) / [(D - N) + 2(N - N_a)] \quad (3)$$

shows nodulated plants to have the value of 0.64 ± 0.14 and non-nodulated plants the value 0.72 ± 0.14 , the latter being very close to the expected value of 0.74.

I conclude from use of the method of CO₂ exchange described here that the energy requirement for symbiotic nitrogen fixation by subterranean clover is substantially greater than that required for the assimilation of mineral nitrogen from the soil.

J. H. SILSBURY

Department of Agronomy,
Waite Agricultural Research Institute,
University of Adelaide,
Glen Osmond, South Australia, 5064

Received 21 December 1976; accepted 18 February 1977.

¹ Gibson, A. H. *Somiplan Symp.* Kuala Lumpur (1976).

² McCree, K. J. in *Prediction and Measurement of Photosynthetic Productivity* (ed. Setlik, I.), 221–229 (Pudoc, Wageningen, 1970).

³ Thornley, J. H. M. *Mathematical Models in Plant Physiology* (Academic, New York, 1976).

⁴ Pirt, S. J. *Proc. R. Soc. B163*, 224–231 (1965).

⁵ Penning de Vries, F. W. T. in *Crop Processes in Controlled Environments* (ed. Rees et al.) (Academic, New York, 1972).

Krameria, free fatty acids and oil-collecting bees

VOGEL'S report^{1,2} of the production of glycerides rather than nectar as rewards gathered by bees described an association involving several genera of insects and numerous plant families. We report here that members of the Krameriaceae produce floral lips more unusual than those described by Vogel². Our data also provide evidence that the oils of *Krameria* are collected by bees of only one genus, *Centris* (Anthophoridae), and used, mixed with pollen, as larval food.

The Krameriaceae is a monotypic family of about 17 species native to the New World arid and semi-arid regions from Kansas, USA to northern Argentina and Chile. The flowers of all species have five (or four) rather showy sepals (Fig. 1a, s) in a zygomorphic arrangement reminiscent of a caesalpinaceae legume flower. The true petals are dimorphic. Three (or two) of the petals are small, but petaloid, often fused at the base, and form a 'flag' above the ovary (Fig. 1a, p). The remaining two petals are highly modified into fleshy secretory glands, one on either side of the ovary (Fig. 1a, g). Two externally distinctive kinds of glands occur in two species groups within the family. One type (Fig. 1b) has bulbous swellings over the densely secreting areas of the lip and upper portion of the gland. The second type (Fig. 1c) secretes oil over the entire pebbled outer surface. Glands of both types produce the oils through a row of elongate epidermal cells (Fig. 1d). The oils collect under the cuticle, which, separates from the epidermal cells, forming lipid-filled 'blisters'.

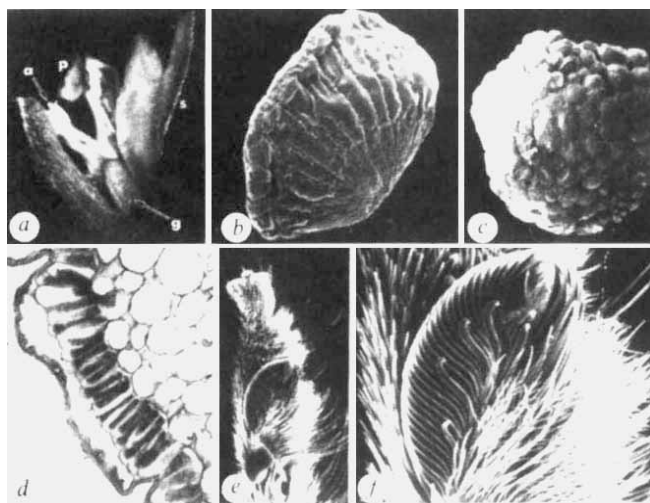


Fig. 1 A *Krameria* flower, oil-producing glands and scraping devices of oil-collecting *Centris*. a, Flower of *K. revoluta*: a, anthers; p, petals; s, sepals; g, gland ($\times 4$). b, Gland of *K. revoluta*, freeze dried, photographed under scanning electron microscopy, (SEM) ($\times 14$). c, Gland of *K. grayi*, critical point dried, photographed under SEM. Where the cuticle has been scraped away, the epidermal secretory cells are visible ($\times 14$). d, Cross section of the edge of a gland of *K. revoluta* showing the secretory epidermal cells and the separated cuticle, light microscopy ($\times 180$). e, Left front leg terminal segment of a female *Centris* (*Paracentris*) *aterrima*, an oil-collecting visitor of *K. revoluta*, under SEM ($\times 18$). f, Basitarsal scraper of the leg in (e) showing the opposing rows of setae, under SEM ($\times 55$).

None of the species in the 49 genera in six plant families² known or suspected to produce floral oils as rewards secretes nectar. In most of these taxa lipids are secreted by trichomes, but members of *Sigmatostalix* and *Oncidium* (Orchidaceae) and all the oil-producing Malpighiaceae have epidermal secreting cells similar to those of *Krameria*.

The only detailed chemical investigations of floral oils have been carried out by Vogel and collaborators² on trichome exudates of *Calceolaria pavonii*. They identified the major component of the crude exudate as a diglyceride of acetic acid and β -acetoxypalmitic acid.

We have analysed the floral lipids of several species of *Krameria* using a combination of thin-layer and gas chromatography, infrared spectroscopy, nuclear magnetic resonance (NMR) and mass spectroscopy. The crude oil eluted from glands reached 0.91 mg per gland. The extracts seemed to be quite pure and similar in all species sampled. Spectral data and thin-layer chromatography after saponification revealed no traces of glycerol and indicated that the major components were free fatty acids. Infrared and NMR spectra showed that the fatty acids were saturated and had an acetate moiety in the β position. Analyses of the methyl esters using gas chromatography indicated chain lengths between C16 and C20. Free fatty acids are rare in plants; as far as we can determine, β -acetate-substituted free fatty acids have not been reported before.

With two exceptions³, all bees presumed to collect floral oils belong to seven New World genera of the Anthophoridae. Our data (Table 1) indicate that only female bees of the genus *Centris* visit *Krameria* for oil collection. The basitarsi of the fore and midlegs of almost all oil-collecting female bees have modifications which facilitate the collection of oils. The size, shape, fine structure and orientation of the oil-collecting scrapers show considerable interspecific and intraspecific variation related in part to the type of oil-producing flower visited (that is those with epidermal cell glands compared with trichomes). All female *Centris* known to visit *Krameria* have a forebasitarsal scraping device (Fig. 1e) consisting of an inner apical comb

Table 1 *Krameria*-*Centris* associations

<i>Krameria</i> species	Locality	<i>Centris</i> species (females only)
<i>K. cistoides</i> Hook & Arn.	Chile	<i>C. (Paracentris) chilensis</i> (Spinola)
	Chile	<i>C. (Wagenknechtia) rhodophthalma</i> Perez
<i>K. cuspidata</i> Presl.	Mexico	<i>C. (Paracentris) aterrima</i> Smith
	Mexico	<i>C. (Paracentris) atripes</i> Mocsary
	Mexico	<i>C. (Hemisiella) trigonoides</i> Lepeletier
	Honduras	<i>C. (Centris) segregata</i> Crawford
<i>K. cytisoides</i> Cav.	Mexico	<i>C. (Acritocentris) n. sp.</i> (Snelling det.)
<i>K. grayi</i> Rose & Painter	Arizona	<i>C. (Paracentris) rhodopus</i> Cockerell
	Arizona	<i>C. (Paracentris) cockerelli resoluta</i> Cockerell
<i>K. lanceolata</i> Torrey	Texas	<i>C. (Paracentris) atripes</i> Mocsary
<i>K. parvifolia</i> Benth.	Arizona	<i>C. (Paracentris) aterrima</i> Smith
	Arizona	<i>C. (Paracentris) rhodopus</i> Cockerell
	Arizona	<i>C. (Paracentris) cockerelli resoluta</i> Cockerell
<i>K. ramosissima</i> (A. Gray) S. Wats.	Texas	<i>C. (Paracentris) atripes</i> Mocsary
<i>K. revoluta</i> Berg.	Mexico	<i>C. (Paracentris) aterrima</i> Smith
	Mexico	<i>C. (Paracentris) atripes</i> Mocsary

of flattened setae with a short opposing row of giant spatulate setae (Fig. 1f). There is also a less elaborate row of flattened setae on the midbasitarsi.

Our field observations of *Centris* females show that their oil-foraging behaviour is quite stereotyped. When approaching a *Krameria* flower, the female orients herself with the supra-ovarian petals and alights, straddling the stamens and stigma. She then grasps the base of these petals with her mandibles and rapidly rubs the scrapers on the forelegs over the glands, rupturing the blisters and accumulating the oil in the hollows of the scrapers. If manipulating oil glands, a female normally spends 2–5 s per flower, but she often approaches and rejects 30 or more flowers in succession before alighting. After accumulating oil from several flowers, she transfers it, while hovering or perched on a twig, to the scopal hairs of the hind legs. The oils can be held alone by capillary action or mixed with pollen for transport to the nest.

The oils seem to be used solely as part of the larval nest provisions. Like most members of the Anthophoridae, members of the genus *Centris* are solitary bees with each female building and provisioning her own nest cells^{3,4}. Each urn-shaped larval cell is lined with a wax-like substance and half filled with a mixture of semifluid pollen and oil, on top of which the egg is laid. Our analyses of the contents of a complete nest cell of *C. trigonoides* from Pochutla, Mexico, showed no nectar (no appreciable sugar) and large quantities of lipids mixed with pollen.

Most species of *Centris* are catholic in their choice of pollen hosts, but they do not specifically collect *Krameria* pollen for larval provision. Nevertheless, they are clearly the primary cross-pollination agents for all *Krameria* species studied. The anthers of *Krameria* flowers (Fig. 1a, a) are clustered around the slightly longer style. When the bee alights, her body first contacts the stigma and then the masses of pollen extruding from the terminal anther pores. Some of this pollen is deposited on the underside of her body (usually on the hairs of the ventral portion of the thorax, the basal portions of the mid and hind legs or on a medial tuft of hairs on the second abdominal sternite) from which it can easily be transferred to the stigma of the next flower visited. The high floral constancy observed among all oil-foraging *Centris* presumably contributes to their efficiency as pollinators. *Krameria* species which we have tested seem to be self-incompatible and require cross-pollination.

The utilisation of a diet very rich in lipids by the larvae of a group of insects generally considered to depend on nectar and pollen (a high carbohydrate, low lipid diet^{5,6}) underscores the ignorance of the nutritional requirements and metabolic capabilities of the most important single group of angiosperm pollinators, the bees. Studies such as ours on oil-producing plants and their pollinators and the chemical nature of the oils may thus reveal not only novel

kinds of lipids but new hymenopteran metabolic pathways as well.

BERYL B. SIMPSON

Department of Botany,
Smithsonian Institution,
Washington DC 20560

JOHN L. NEFF

Division of Natural Sciences,
University of California,
Santa Cruz, California 95064

DAVID SEIGLER

Department of Botany,
University of Illinois,
Urbana, Illinois 61801

Received 14 December 1976; accepted 15 March 1977.

¹ Vogel, S. *Abstr. ninth int. bot. Congr.*, Seattle 299 (1969).

² Vogel, S. *Naturwissenschaften* 58, 58 (1971); *ibid* 63, 44 (1976); *Olblumen und oltsammelnde Bienen* (Steiner, Wiesbaden, 1974); *Fifteenth int. ent. Congr.*, Washington, D. C., Aug 19–27 (1976).

³ Michener, C. D. & Lange, R. B. *Univ. Kansas Sci. Bull.* 39, 69–96 (1958).

⁴ Silvestri, F. *Redia* 1, 1–234 (1903).

⁵ Baker, H. G. & Baker, I. in *Coevolution of Animals and Plants* (ed. Gilbert, L. & Raven, P. H.) 100–140 (University of Texas, Austin, 1975).

⁶ Stanley, R. G. & Linskins, H. F. (eds) *Pollen: Biology, Biochemistry, Management* (Springer, New York, 1974); Lunden, R. *Grana Palynol.* 1 (2), 3–19 (1956).

Mating advantage of rare males in models of sexual selection

MODELS of sexual selection in polygynous species of animals have been derived on the assumption that some females have preferences to mate with males with particular genotypes^{1–3}. The mating advantage gained by the males is always frequency-dependent because the preferred males take part in the same number of preferential matings when they are rare as when they are common; individually therefore, they mate more often when they are rare. Frequency-dependent sexual selection has been demonstrated in many experiments with *Drosophila*: rare males take part in a higher proportion of matings than their frequency as available mates^{4–12}. Ehrman and Spiess¹⁰ explained this phenomenon by frequency-dependence either in female preference or in male courtship. This explanation, which is difficult to interpret in behavioural terms, may not be necessary, however, because constant female preferences would entail frequency-dependent selection among the males. I show here that a simple model of constant preferences for particular phenotypes or genotypes is sufficient to explain a large body of data on frequency-dependent sexual selection in *Drosophila*.

In models analysed previously^{1–3}, some females had mating preferences either for particular genotypes or for a phenotype determined by a dominant allele; other females mated at random. These possibilities can be combined in a model that allows for any degree of dominance as well as separate

Table 1 Fitting data of Spiess and Spiess to the model with constant mating preferences

(a) Matings with karyotypes homozygous for inversions AR and PP in <i>Drosophila pseudoobscura</i> ⁸						
Before fitting the model $\chi^2 = 67.105$						
Maximum likelihood estimates: preference for AR, $\alpha = 0.138$						
preference for PP, $\theta = 0.021$						
Ratio among males		Observed nos of matings of males		Expected nos of matings of males		Contribution to χ^2 after fitting
AR	PP	AR	PP	AR	PP	
2	18	51	214	58.856	206.144	1.348
4	16	63	135	60.627	137.373	0.134
6	14	38	60	38.249	59.751	0.003
8	12	56	43	46.965	52.035	3.307
10	10	49	37	48.031	37.969	0.044
12	8	69	37	68.115	37.885	0.032
14	6	78	19	70.489	26.511	2.928
16	4	182	41	180.807	42.193	0.042
18	2	292	40	297.104	34.896	0.834
After fitting the model $\chi^2 = 8.672$						
This value of χ^2 for 7 degrees of freedom shows there is no residual heterogeneity after the fitting of the model, the two parameters contributing $\chi^2 = 58.433$ to the overall χ^2 .						
(b) Matings with strains Humboldt (Hu) and white wolf (WW) in <i>Drosophila persimilis</i> ⁹						
Before fitting the model $\chi^2 = 101.707$						
Maximum likelihood estimates: preference for Hu $\alpha = 0.166$						
preference for WW $\theta = 0.156$						
Ratio among males		Observed nos of matings of males		Expected nos of matings of males		Contribution to χ^2 after fitting
Hu	WW	Hu	WW	Hu	WW	
2	18	53	194	57.656	189.344	0.491
4	16	25	34	17.776	41.224	4.201
6	14	17	28	16.612	28.388	0.014
8	12	18	17	15.296	19.704	0.849
10	10	25	25	25.245	24.755	0.005
12	8	15	20	20.047	14.953	2.974
14	6	24	13	23.704	13.296	0.010
16	4	32	23	38.968	16.032	4.275
18	2	175	43	169.251	48.749	0.873
After fitting the model $\chi^2 = 13.692$						
This value of χ^2 shows no residual heterogeneity, the two parameters contributing $\chi^2 = 88.015$ to the overall χ^2 .						

In the data of this table the matings of the two male genotypes are recorded in relation to the genotypes of females which were in a constant ratio of 50% while the males' frequencies were varied. Although PP and WW females mated less frequently than AR and Hu females, there is no evidence for assortative mating preferences in the two experiments and no heterogeneity in the relative frequencies of the matings of the males with the two female genotypes. In this table, therefore, the numbers of matings with the two genotypes of females have been added together.

preferences for each genotype. The preferences can be defined as follows:

genotypes	AA	AA'	A'A'
frequencies of genotype in population	u	v	w
proportions of females preferring each genotype separately	α'	β'	γ'
proportions of females preferring 'A' phenotypes	$au/(u+v)$	$av/(u+v)$	—

It can then be shown¹³ that in the following generation the genotypic frequencies become

$$\begin{aligned} u' &= \phi p + \alpha p^2/(1-w) + p^2(1-\alpha-\theta) \\ v' &= \phi q + (\theta-\phi)p + \alpha p(2q-w)/(1-w) + 2pq(1-\alpha-\theta) \\ w' &= (\theta-\phi)q + \alpha q(q-w)/(1-w) + q^2(1-\alpha-\theta) \end{aligned}$$

where $p = u + \frac{1}{2}v$ and $q = w + \frac{1}{2}v$ are the gene frequencies of A and A' ; and $\phi = \alpha' + \frac{1}{2}\beta'$, $\theta = \alpha' + \beta' + \gamma'$. Thus $\alpha + \theta$ is the total proportion of all females who mate preferentially; the remaining proportion $1 - \alpha - \theta$ mate at random. A stable polymorphic equilibrium is established at a gene frequency of A'

$$q_* = \frac{-\phi + \sqrt{(\phi^2 + 4(\theta-\phi)(\alpha+\theta))}}{2(\alpha+\theta)} \quad (1)$$

For a change of gene frequency from q_0 to q_n over a period of n generations, approximately

$$\begin{aligned} \left\{ \frac{q_n - q_*}{q_0 - q_*} \right\}^{(q_*+1)} & \left\{ \frac{q_n + q_* + a}{q_0 + q_* + a} \right\}^{(q_*+a-1)} \\ &= \exp[-n(q_*+a)(\alpha+\theta)] \end{aligned} \quad (2)$$

where $a = \phi/(\alpha+\theta)$.

Two cases are likely to apply to sexual selection in natural populations. When $\phi = \theta = \alpha'$ and $\beta' = \gamma' = 0$, A is at an unconditional advantage and spreads through the population to complete fixation. Among the females, a proportion α are more responsive to both AA and AA' males and a proportion ϕ are more responsive only to AA males. Thus in its effect on mating preference A is being selected as a semi-dominant or semi-recessive allele depending on the magnitudes of ϕ and α . In addition, some females may prefer $A'A'$ males. Then $\phi = \alpha'$ and $\theta = \alpha' + \gamma'$ where γ' is the proportion of females preferring $A'A'$ males. This produces a stable polymorphism at an equilibrium frequency given by equation (1). When A is completely dominant $\phi = \alpha' = 0$ and

$$q_* = \sqrt{(\theta/(\alpha+\theta))} \quad (3)$$

In this case $a = 0$ and the change in gene frequency over a period of n generations becomes

$$\begin{aligned} \left\{ \frac{q_n - q_*}{q_0 - q_*} \right\}^{(q_*+1)} & \left\{ \frac{q_n + q_*}{q_0 + q_*} \right\}^{(q_*-1)} \\ &= \exp[-nq_*(\alpha+\theta)] \end{aligned} \quad (4)$$

These models explain the establishment of stable polymorphisms in population cages of *Drosophila*. Equation (4) could be used to estimate the values of q_* and the total preference, $\alpha + \theta$. Combined with direct estimates of α and θ from observations of the numbers of the different matings, this would provide an independent test of the model.

If observations have been made of the numbers of matings for a range of different frequencies, the estimates of α and θ

can also be tested for heterogeneity against the null hypothesis that the proportions of females with preferences, α and θ , remain constant and determine the numbers of matings at each different frequency of the male genotypes. This test can be applied to Ehrman's and Spiess' data of number of matings of males at varying frequencies. Males of two different genotypes were used in the experiments. Females were introduced into the mating chamber containing the males and the genotype of the mated male was recorded. According to the model, the expectations are as follows:

genotypes	<i>AA</i>	<i>BB</i>
frequencies of genotypes		
of males	<i>u</i>	<i>v</i>
proportions of females		
preferring each genotype	α	θ
proportions of matings		
with each genotype	$\alpha + u(1 - \alpha - \theta)$	$\theta + v(1 - \alpha - \theta)$

Suppose, in the i th experiment a_i and b_i are the numbers of matings observed with A and B males at frequencies u_i and $v_i = 1 - u_i$, then the log likelihood for all of k experiments is

$$\log_e L = \sum_{i=1}^k a_i \log[\alpha + u_i (1 - \alpha - \theta)] + \sum_{i=1}^k b_i \log[\theta + v_i (1 - \alpha - \theta)]$$

Maximum likelihood estimates of α and θ are then obtained as the values which maximise $\log_e L$. I have calculated these estimates by trial and error, starting with arbitrarily chosen values of α and θ and introducing random perturbations in the initial values until higher values of $\log_e L$ are found. By making the maximum size of the random perturbations smaller and smaller, the maximum of the log likelihood surface is reached by a random walk of successively smaller steps.

In Ehrman's experiments^{5, 6, 11}, the frequency of the female genotypes varied with the frequency of the male genotypes. Spiess^{8, 9}, however, kept the frequencies constant among the individuals who were being offered the choice of mates. Choice was generally exercised only by the females, the males mating at random when offered a choice of female genotypes. In his experiments on female choice, Spiess varied the frequencies of the male genotypes between experiments while maintaining the frequencies of the genotypes at 50% in the females. Thus, if choice varied according to genotype, this could not be a cause of the frequency-dependent advantage gained by the males in Spiess' experiments. Table 1 shows the data Spiess obtained and the fit of the model. The model provides a satisfactory fit with no residual heterogeneity after the expected numbers of matings have been calculated using the maximum likelihood estimates of α and θ from the total data of each of the experiments.

Table 2 shows the results of fitting Ehrman's data⁵ on the effect of frequency on the matings of males homozygous for the karyotypes CH and AR in *Drosophila pseudoobscura*. In these

Table 2 Fitting data of Ehrman⁵ on matings with karyotypes homozygous for inversions CH and AR in *Drosophila pseudoobscura* to the model with constant mating preferences

(a) Matings in a line that had been selected for positive geotaxis in previous generations								
All matings					Early matings			
Ratio among males		Contribution to χ^2 after fitting			Observed nos of matings		Contribution to χ^2 after fitting	
CH	AR	CH	AR		CH	AR		
18	2	60	50	3.105	31	25		1.816
16	4	66	51	0.140	36	24		0.025
14	6	72	48	2.000	37	25		0.809
12	8	44	63	2.662	21	33		2.198
10	10	51	53	0.887	23	30		0.007
8	12	57	47	9.645	31	23		7.731
6	14	47	60	3.452	26	28		4.857
4	16	35	83	0.070	14	45		0.778
2	18	15	93	8.499	7	48		3.791
Preference for CH		$\alpha = 0.217$					$\alpha = 0.189$	
Preference for AR		$\theta = 0.328$					$\theta = 0.310$	
Deviation from								
$\alpha = \theta = 0$		$\chi^2_2 = 213.591$					$\chi^2_2 = 98.014$	
Residual heterogeneity		$\chi^2_7 = 30.460$					$\chi^2_7 = 22.013$	

There is significant residual heterogeneity for the matings with male ratios 8:12 and 2:18. There are significantly too many matings of the less common males at the ratio 8:12 and too few matings of the less common males at the ratio 2:18. Otherwise, the fit of the model is very good with no consistent deviations from the model.

(b) Matings in a line that had been selected for negative geotaxis in previous generations								
All matings					Early matings			
Ratio among males		Observed nos of matings		Contribution to χ^2 after fitting	Observed nos of matings		Contribution to χ^2 after fitting	
CH	AR	CH	AR		CH	AR		
18	2	78	27	3.089	45	9		2.728
16	4	75	43	0.208	39	20		0.062
14	6	63	46	0.039	36	20		0.143
12	8	42	75	12.461	24	36		6.268
10	10	49	66	1.122	26	32		0.683
8	12	57	63	1.044	28	35		0.000
6	14	37	71	0.720	21	35		0.033
4	16	46	64	3.356	24	32		2.517
2	18	33	80	0.005	16	43		0.000
Preference for CH				$\alpha = 0.243$			$\alpha = 0.213$	
Preference for AR				$\theta = 0.292$			$\theta = 0.208$	
Deviation from								
$\alpha = \theta = 0$								
Residual heterogeneity				$\chi^2_2 = 148.036$			$\chi^2_2 = 48.361$	
				$\chi^2_7 = 22.044$			$\chi^2_7 = 12.435$	

There is still some residual heterogeneity among all matings after the fit of the model. This heterogeneity is determined by the high frequency of matings of the less common males at the ratio of 12:8.

experiments, data at some frequencies show significant deviations from expectation after the fitting of the model: the ratios of 8:12 in the positively selected line and 12:8 in the negatively selected line show a great excess of matings of the less common males. No simple model would resolve these anomalies because at these ratios the mating advantage shows no particular relationship, frequency-dependent or not, with the mating advantages at the other ratios: there are no consistent deviations from the model. If preferences were genetically determined, chance would produce sampling variation in the proportions of females expressing a preference. If α and θ do vary between experiments, then the variance in the observed numbers of matings will be greater than the binomial variance. This will therefore inflate the values of χ^2 which have been calculated on the assumption of binomial variation. I consider this to be the most likely cause of the statistically significant heterogeneity in Ehrman's data. A model of the expected variance in α and θ would have to be constructed in order to calculate a corrected value of χ^2 .

Ehrman's experiment¹¹ on frequency-dependent mating for two recessive alleles *or* (orange) and *pr* (purple) also shows residual heterogeneity that is just statistically significant, and also without any systematic deviations from the model. In this experiment population cages were started at ratios of 80 *or*: 20 *pr* and 20 *or*: 80 *pr* and continued for 10 generations. In each generation the actual matings of 100 pairs were observed. It is to be expected that the mating preferences would vary between generations mainly as a result of sampling genetic drift. The variance in the observed numbers of matings must be greater than the binomial and χ^2 will therefore be inflated. Fitting the model to the data gives $\alpha = 0.244$ for the preference for *or* and $\theta = 0.259$ for the preference for *pr*. Before fitting the parameters, $\chi^2_{20} = 64.132$; and after fitting, $\chi^2_{18} = 30.823$ ($0.05 > P > 0.02$).

Ehrman⁷ tried to show "how a *Drosophila* female knows which males are rare". On my theory, the females do not have to carry out the difficult task of discriminating between rare and common males: constant mating preferences are sufficient to explain the observed frequency-dependence. In some of Ehrman's experiments, the mating advantage disappeared if the rare males in the mating chamber were the common males in an adjacent chamber from which air was passed into the mating chamber. This was interpreted to mean that olfactory stimuli from the adjacent chamber where the males were common prevented the discrimination of rare males in the mating chamber. The results, however, were "inconsistent and ambiguous" in some of these experiments. Sampling variation in the proportions of females expressing preferences might be the cause of both this inconsistency and the residual heterogeneity after fitting the data of some of Ehrman's experiments to my model.

O'Donald¹³ discussed the possible behavioural mechanisms which might give rise to mating preference. Some females may have a lower threshold of response to some male genotypes and would require less stimulation to mate with them. They would thus mate more readily with the preferred males and more readily than the other females who mate at random. The presence of other courting males will raise the general level of stimulation so that all females are stimulated to mate more readily. A female with a lower threshold towards a particular genotype of male would then mate more readily with any male. The proportion of preferential matings would therefore decline with increasing density. The mating preference, according to this model, will be density-dependent but not frequency-dependent. The disappearance of mating choice in some of Ehrman's experiments could be the results of the increased level of stimulation provided by the extra males in the adjacent chamber. A similar result would follow if the different male genotypes determined different levels of courtship activity. A more responsive female would be sufficiently stimulated by any male and so would mate more or less at random. A less responsive female would mate more readily with one of the more active males and would therefore tend to show a preference

that decreased at higher levels of stimulation. The mating preferences could thus be determined by variation either in male courtship or in female response. Both types of behavioural mechanism would explain observations of mating choice in *Drosophila*.

PETER O'DONALD

Department of Genetics,
University of Cambridge,
Cambridge, UK

Received 17 January; accepted 17 March 1977.

- ¹ O'Donald, P. *Heredity* **18**, 451 (1963).
- ² O'Donald, P. *Heredity* **22**, 499 (1967).
- ³ O'Donald, P. *Heredity* **31**, 145 (1973).
- ⁴ Ehrman, L. *Anim. Behav.* **14**, 332 (1966).
- ⁵ Ehrman, L. *Am. Nat.* **101**, 415 (1967).
- ⁶ Ehrman, L. *Genet. Res., Camb.* **11**, 135 (1968).
- ⁷ Ehrman, L. *Evolution* **23**, 59 (1969).
- ⁸ Spiess, E. B. *Am. Nat.* **102**, 363 (1968).
- ⁹ Spiess, L. D. & Spiess, E. B. *Am. Nat.* **103**, 155 (1969).
- ¹⁰ Ehrman, L. & Spiess, E. B. *Am. Nat.* **103**, 675 (1969).
- ¹¹ Ehrman, L. *Proc. natn. Acad. Sci. U.S.A.* **65**, 345 (1970).
- ¹² Ehrman, L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) (Aldine, Chicago, 1972).
- ¹³ O'Donald, P. *J. theor. pop. Biol.* (in the press).

Spontaneous forward mutation versus reversion frequencies for maize *Adh1* in pollen

RELIABLE quantitative data on spontaneous, specific gene mutation frequencies in higher plants and animals are few. Detergents include low natural frequencies and difficulty in obtaining in excess of 10^6 scorable organisms or gametophytes. The *alcohol dehydrogenase-1* gene (*Adh1* gene; ADH enzyme EC 1.1.1.1.), as expressed in pollen grains, is among the exceptionally suitable; much is known about the maize ADHs¹⁻³, *Adh1* function is totally dispensable in an aerobic environment and maize pollen is a trinucleate gametophyte which expresses much of its haploid genome, including *Adh1*^{4,5}. In particular, there have been two recent methodological advances. First, I am able to cytochemically stain pollen, before⁶ or after *in vitro* germination, for the presence of above 5% normal ADH activity. And second, ADH⁻ pollen grains survive allyl alcohol (C=C-C-OH) vapour concentrations which kill ADH⁺ grains^{6,7}; this selection scheme was developed for yeast by Megnet⁸. My genetic resolution is approximately one mutant (*Adh1*⁺→ADH⁻) per 10^7 chemically selected, viable gametophytes, and one (phenotypic) revertant (*Adh1*⁻→ADH⁺) per 10^8 unselected gametophytes. In this note, I compare spontaneous forward mutant frequency with previously published⁹ revertant frequencies for one naturally occurring and six ethyl methanesulphonate-induced *Adh1*-deficient (*Adh1*⁻) alleles.

About 2% of normal, *Adh1*⁺ pollen grains stain poorly in the ADH-specific reagents, a condition which almost always denotes inviability. For that reason, it was necessary to count ADH⁻ gametophytes among viable pollen rather than among total pollen. My methods were developed by using *Adh1*⁺/*Adh1*⁻ (*Adh1*- γ 25; CRM⁻; unpublished) heterozygotic plants segregating 1:1 for blue, opaque: yellow, translucent pollen. Pollen was collected under controlled conditions, cold treated, exposed to a series of allyl alcohol vapour concentrations, germinated on a slice from 'David's Bread Loaf' medium and subsequently frozen. The ice-perforated grains were then stained for ADH activity (methods described in Fig. 1 legend). Figure 1 shows the final result for an allyl alcohol concentration which kills 99.999% ADH⁺ pollen without affecting the viability of the 50% ADH⁻ grains: none of the blue, opaque grains have pollen tubes.

Wild-type *Adh1*⁺ pollen (either *Adh1-S* or *Adh1-F* isoalleles) was collected in an isolation greenhouse and treated according to the protocol outlined above. Allyl alcohol

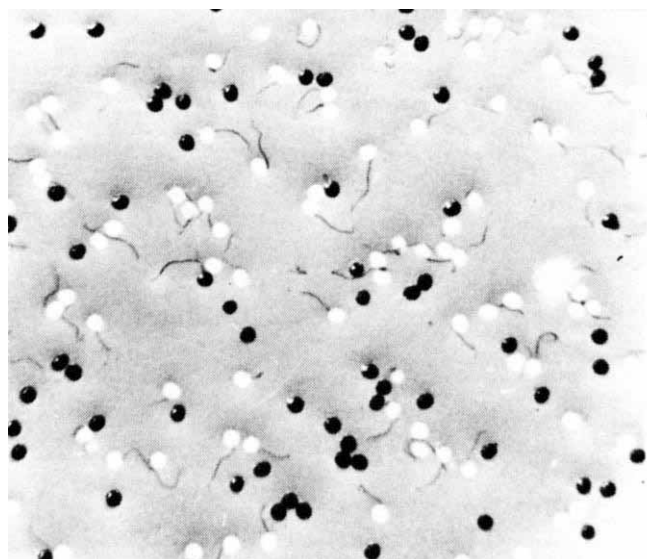


Fig. 1 Pollen shed from an *Adhl*⁺/*Adhl*⁻ plant after allyl alcohol treatment, *in vitro* germination, and *in situ* staining for ADH activity. Pollen was collected before 10 a.m., weighed, stored desiccated (CaSO₄) at 4 °C for 3 h and treated with 300 μ M allyl alcohol for 30 min in a 500 ml sealed Mason jar. The treated pollen was brushed on to 2 mm slices from 'David's Bread Loaf' medium⁷, much modified from that of Cook and Walden¹¹, and allowed to germinate for 45 min. The slabs were frozen at -23 °C for 3 h and defrosted 30 min at 25 °C. ADH-specific stain was layered over the gametophytes three times, each for 30 min as previously described⁵. ADH⁺ gametophytes stained blue and opaque.

concentrations were chosen which allow a few ADH⁺ gametophytes to germinate. A forward mutant (*Adhl*⁺→ADH⁻) would be a yellow, translucent pollen grain with tube (allyl alcohol resistant) surrounded by blue, opaque non-germinated grains. Since pollen is naploid, ADH⁻ mutants might include modifiers. As shown in Table 1,

there were zero viable ADH⁻ grains among 4.9×10^6 viable gametophytes. The reliability of these data is supported by the following two considerations. (1) That allyl alcohol does not appreciably affect the viability of ADH⁻ pollen grains was insured by using borderline concentrations where some ADH⁺ gametophytes germinate; previous experiments⁷ indicated that the allyl alcohol concentrations I used killed no more than 20% ADH⁻ grains. (2) As a further control, I scored pollen shed from wild-type plants after the immature tassel containing the pollen mother cells was X-irradiated (last line, Table 1). Since the radiation-induced mutant frequency was rather high, these methods would have detected ADH⁻ gametophytes if they had occurred (*Adhl*⁺→ADH⁻ in mutants/viable gametophyte) at a frequency much above 2×10^{-7} .

My previous work⁵ quantified the (phenotypic) revertant frequencies (ADH⁺ revertants/unselected gametophyte) for one naturally occurring and six ethylmethane sulphonate-induced *Adhl*⁻ mutants. These mutants, recovered and characterised by Schwartz (see ref. 5) were chosen at random; two are CRM⁻ while the remaining five have lesions involving the structural gene component of the cistron. Five mutants are from the same inbred line⁸. A revertant⁹ includes all mutants capable of changing an *Adhl*⁻ (yellow, translucent) to an ADH⁺ grain (blue; opaque): true revertants, second site revertants, unlinked suppressors; supersuppressors and overproducers are also possible for some mutants. An interesting mode of reversion would be the 'derepression' of the *Adh2* gene which is generally switched off in pollen; although possible, this explanation for high revertant frequency is not likely⁵. Reversion involving altered metabolic circuitry is not a factor since I measure *in situ* enzyme activity directly. The revertant frequencies ± 1 s.e. $\times 10^6$ were 1.6 ± 0.4 ; 7.3 ± 1.4 ; 2.1 ± 1.0 ; 7.2 ± 1.1 ; 10.8 ± 1.7 ; 1.2 ± 0.3 and 5.7 in ADH⁺ per gametophyte. The median revertant frequency was 5.7×10^{-6} ADH⁺ per gametophyte. This frequency is an underestimate, since 5% normal ADH activity and perhaps as

Table 1 Spontaneous *Adhl*⁺ to ADH⁻ mutant frequency

F ₂ <i>Adhl</i> Genotype	Control % Germination	Grains $\times 10^{-5}$ (Total)*	(Viable)	C- C-C-OH (μ M)	No. (or %) germination† Post-selection†	No. ADH ⁻
S/F (12 plants)	76.25	4.00	3.05	222	0	0
	72.50	4.00	2.90	222	0	0
	92.50	4.00	3.70	222	0	0
	83.00	5.00	4.15	222	0	0
	70‡	4.00	2.80‡	37	0	0
	70‡	3.00	2.10‡	37	0	0
	76.00	3.60	2.74	15	23	0
	90.00	1.60	1.44	15	4	0
	86.25	2.00	1.73	15	18	0
	84.50	3.40	2.87	6	(41.5)	0
	86.30	2.92	2.52	15	(32.0)	0
	88.00	3.00	2.64	15	(16.3)	0
F/F (4 plants)	79.00	4.00	3.16	222	0	0
	70‡	2.14	1.50‡	37	0	0
	70.00	4.00	2.80	37	0	0
	76.7	3.00	2.30	7§	(12.0)	0
S/S (3 plants)	70‡	2.55	1.78‡	37	0	0
	93.60	1.60	1.50	15	(40.0)	0
	82.50	3.72	3.07	6§	(73.5)	0
		Total	48.75			0
S/S (5 plants) (X-rayed)	45.0 (incl. 40% abortion)	6.64	2.99	148	57	19 (or 6.4×10^{-5} mutants per viable gametophyte)

* No. grains calculated on the basis of control experiments indicating that 1 mg fresh pollen of these families equals 2,000 grains. The X-irradiated plants gave 40% aborted pollen; we assumed that weight of an abortant was 10% that of a plump grain.

† Every allyl alcohol survivor was stained in hopes of finding an ADH⁻.

‡ Minimum values.

§ Treated for 25 rather than 30 min.

|| 720 rad at 56 rad min⁻¹ were delivered to the pollen mother cells. Undoubtedly, some of the germ cells were premeiotic while others were past microspore mitoses; these data are not to be quantified.

Table 2 Paired data on spontaneous forward mutation versus median reversion frequencies for specific genes

Organism	Cistron(s)	<i>f</i> (forward mutation)	Units	Ref.	Median <i>f</i> (reversion)	Units	No. alleles	Ref.	$\frac{f \text{ (reversion)}}{f \text{ (forward mutation)}}$
T-4	<i>RIIA</i> and <i>RIIB</i>	4.5×10^{-4}	75% r/colony	12	$5 \times 10^{-8*}$	revertants/viable phage	20	13	10^{-4}
<i>N. crassa</i>	<i>Ad-3A</i> and <i>Ad-3B</i>	3.8×10^{-7}	mutants/viable conidium	14	$2 \times 10^{-6*}$	revertants/viable conidium	4†	15	5.3
Maize	<i>wx</i>	$<7 \times 10^{-7}$	mutants/seed	16	7×10^{-6}	revertants/pollen	14	17	>10
	<i>Adhl</i>	$<2 \times 10^{-7}$	mutants/viable pollen	—	5.7×10^{-8}	revertants/pollen	7	—	>29
Mouse	<i>a, b, c, d,</i> and <i>In</i> pooled	$1.3 \times 10^{-6\dagger}$	mutants/locus/mouse	18	$2.7 \times 10^{-6*}$	revertants/locus/mouse	5§	18	0.2

* All were either true revertants or other intragenic events.

† Of 24 X-ray induced mutants, only these four reverted at frequencies above 2×10^{-8} .

‡ Macrolesions pass through the gametes.

§ One allele for each of these five loci; only *a* and *d* gave revertants where *f*s below 2×10^{-6} per locus per mouse go undetected.

|| Auerbach and Kilbey¹⁸ commented: "Surprisingly, forward and reverse mutation rates at five of these loci differed only by a factor of three."

high as 30% normal ADH activity/gametophyte is scored as ADH⁺ (ref. 5). Recovery and analysis of these revertants is most important and is underway. In any case, $5.7 \times 10^{-9} / <2 \times 10^{-7}$ gives a >29-fold excess of spontaneous revertant over forward mutant frequency.

The comparison of forward mutant and revertant frequencies for specific genes is fraught with methodological pitfalls, hidden assumptions and statistics of dubious applicability. Table 2 compiles some comparative data for a few genes in a few organisms. These data comprise the totality of reliable data found—no selection was intended. Some of the more apparent considerations affecting the reliability of these data follow. Higher animals are expected to show inflated forward mutant frequencies in the absence of rigorous genetic analyses because there is no haploid (gametophytic) stage to impede the transmission of macrolesions. Ideally, a revertant frequency per gene should be based on the median frequency of several revertable alleles. Spontaneous reversion data for *Drosophila* is for 'unstable' alleles or for too few alleles (M. M. Green, personal communication; Green also notes that high reversion is expected of insertion-type mutants). Marker effects perhaps reflecting the mutagenic agent, the contribution of alternative metabolic pathways to reversion frequencies, meiotic effects, variation in modifiers of spontaneous mutagenesis, cell lineage differences between sexes, and differences in germ cell development are all known to affect mutation rates in certain cases. The fundamental nature of the cistron itself might also differ among different organisms^{9,10}. It is generally believed that forward mutant frequencies exceed median revertant frequencies, especially when suppressors are ruled out. But the data for the purple adenine mutants in *Neurospora* and five coat colour loci in mice (Table 2) do not fit this rule. Observations on red adenine mutants in yeast—the homologue of purple adenines in *Neurospora*—support the canonical view that spontaneous forward mutation exceeds median reversion frequencies (H. Roman, personal communication). The clear excess of revertant frequency over forward mutant frequency for the two cistrons studied in maize might be explained readily if almost all revertants were suppressors. However, for both *Adhl* and *wx* (Table 2), every mutant gave relatively high revertant frequencies; nonsense suppression is not likely.

The significance of our *Adhl* comparison and the few other results where spontaneous reversion exceeds forward mutation frequency cannot be assessed readily. It is clear that evolutionary and other arguments resting heavily on

the assumption that reversion frequencies are comparatively negligible should be reevaluated.

This work was supported by the US NIH and the California Agricultural Experiment Station. I thank David S. K. Cheng for expert technical assistance and Roy Sahara for nursery work.

MICHAEL FREELING

Genetics Department,
University of California,
Berkeley, California 94720

Received 17 September 1976; accepted 8 March 1977.

- Schwartz, D. *Proc. natn. Acad. Sci. U.S.A.* **56**, 1431–1436 (1966).
- Felder, M. R., Scandalios, J. G. & Liu, E. H. *Biochim. biophys. Acta* **318**, 149–159 (1973).
- Freeling, M. & Schwartz, D. *Biochem. Genet.* **8**, 27–36 (1973).
- Schwartz, D. *Genetics* **67**, 411–425 (1971).
- Freeling, M. *Genetics* **83**, 701–717 (1976).
- Schwartz, D. & Osterman, J. *Genetics* **83**, 63–65 (1976).
- Cheng, D. S. K. & Freeling, M. *Maize Genet. Coop. News Lett.* **50**, 11–13 (1976).
- Megnet, R. *Archs Biochem. Biophys.* **121**, 194–201 (1967).
- Judd, B. H., Shen, M. W. & Kaufman, T. C. *Genetics* **71**, 139–156 (1972).
- Abrahamson, S., Bender, M. A., Conger, A. D. & Wolff, S. *Nature* **245**, 460–461 (1973).
- Cook, F. & Walden, D. *Can. J. Bot.* **43**, 779–786 (1965).
- Drake, J. W. *J. cell comp. Physiol.* **64**, Suppl. 1, 19–32 (1964).
- Freese, E. *Proc. natn. Acad. Sci. U.S.A.* **45**, 623–633 (1959).
- Brockman, H. E. & de Serres, F. J. *Genetics* **48**, 597–604 (1963).
- de Serres, F. J. *Genetics* **43**, 187–206 (1958).
- Stadler, L. J. *The Spragg Memorial Lectures* (Third Series) 3–15 (Michigan State College, East Lansing, 1942).
- Nelson, O. E. *Genetics* **60**, 607–624 (1968).
- Schlager, G. & Dickie, M. M. *Genetics* **57**, 319–330 (1967).
- Auerbach, C. & Kilbey, B. J. *A. Rev. Genet.* **5**, 163–218 (1971).

Determinant selection is a macrophage dependent immune response gene function

IMMUNE response (Ir) genes are linked to the species histocompatibility complex and define as yet uncharacterised phenotypic products which control the immune response to thymus dependent antigens¹. Antibody formation² and antigen induced T lymphocyte proliferation³ are two examples of immune phenomena which, *in vivo* and *in vitro*, operate under Ir gene influence. To clarify their mechanism of action and cellular location, we have examined the contribution of antigen structure (amino acid sequence and conformation to Ir gene control of antigen recognition by T lymphocytes) as well as to the critical role played by the antigen presenting macrophage in expression of that control^{4–7}. We report that immune response gene control of antigen recognition operates at least in part at the level of the macrophage.

Recent studies of *in vitro* T-cell proliferation in insulin in inbred guinea pigs offered us a unique opportunity to examine, operationally, the site of Ir gene product function⁴. T cells from both strain 2 and 13 guinea pigs immunised with pork insulin in complete Freund's adjuvant proliferate *in vitro* response to insulin. The determinant recognised by the strain 13 guinea pig resides in the B-chain region of the insulin molecule as evidenced by the capacity of insulin immune strain 13 but not strain 2 guinea pigs to respond to isolated oxidised B chain. When strain 2, 13 and F₁ (2×13) guinea pigs are immunised with oxidised B chain, strain 13 and F₁, but not strain 2, respond to either isolated B chain or native insulin. Genetic analysis of backcrossed and phenotypically characterised outbreds immunised to B chain showed that the relevant immune response gene defining response to the B chain determinant is linked to the strain 13 guinea pig histocompatibility complex. By contrast, the strain 2 inbred guinea pig T cell recognised insulin through a determinant present in insulin A chain. More specifically, they recognised the A chain loop and their T-cell response is sensitive to amino acid changes in position A8, A9 and A10. Insulins with total identity in this area completely cross-react in T cell proliferation and T helper cell activity. Insulins with partial or no identity in these amino acids residues show partial or no cross-reactivity respectively. In strain 13 animals, different species variants of insulin show cross-reactivity at a T-cell level independent of their A chain loop constitution. As anticipated, F₁ T cells from insulin immune guinea pigs are able to recognise either determinant.

We assessed the function of macrophages in Ir control taking advantage of the capacity of T cells from insulin immunised F₁ (2×13) guinea pigs to respond to antigen bearing parental macrophages. We reasoned as follows.

If Ir gene function operates exclusively at the level of the responding lymphocyte, both determinants should be simultaneously recognised by F₁ T cells independently of which parental macrophage (Mφ) is 'presenting' the antigen. Conversely, if the definition of which determinant is actually recognised depends on the genetic profile of the antigen presenting Mφ, then Ir gene function must also be operating at this cell level. The data that follow favour the latter assumption. Mφ from both inbred strains of guinea pig are able to 'present' pork insulin to pork insulin immune F₁ (2×13) lymphocytes (Table 1). As with PPD, a strong proliferative response is observed when pork insulin pulsed Mφ from either strain of guinea pig are added to F₁ PELs. When different species variants of intact insulin are presented to the same pork insulin immune F₁ PELs by either strain 2 or 13 Mφ, however, two different patterns of response are seen (Fig. 1). An A-chain loop pattern of cross-reactivity is observed when the antigens are presented on

Table 1 DNA synthetic responses of T cells from pork insulin immunised F₁ (2×13) guinea pigs to parental macrophages pulsed with native insulin

Macrophages pulsed with:	Macrophage genetic background	
	Strain 2	Strain 13
	³ H-TdR incorporation Δc.p.m. × 10 ⁻³	
PPD	197.45 ± 6.60	175.04 ± 29.91
Pork insulin	45.51 ± 11.34	62.42 ± 10.59

F₁ (2×13) guinea pigs were immunised with 100 μg of pork insulin in CFA, 2–3 weeks before use. Oil induced, macrophage-rich peritoneal exudate cells were pulsed with 100 μg ml⁻¹ antigen and 40 μg ml⁻¹ of mitomycin C at 37 °C for 60 min. After four washes 1.2 × 10⁵ of these cells were added to 2.4 × 10⁵ responding peritoneal exudate lymphocytes depleted of adherent cells (PELs) in 200 μl of 5% heat inactivated male guinea pig serum in RPMI-1640 plus 2.5 × 10⁻⁵ M of 2-mercaptoethanol in round microtitre plates. After 48 h of culture 1 μCi of ³H-methyl-thymidine (6.7 Ci mmol⁻¹) was added for an additional 24 h. Cells were collected on glass fibre filters with a semi-automated microharvester (Adaps) and ³H-thymidine incorporation determined by liquid scintillation spectrometry. Data are mean ± s.e.m. from four experiments. PPD, purified protein derivative of tuberculin.

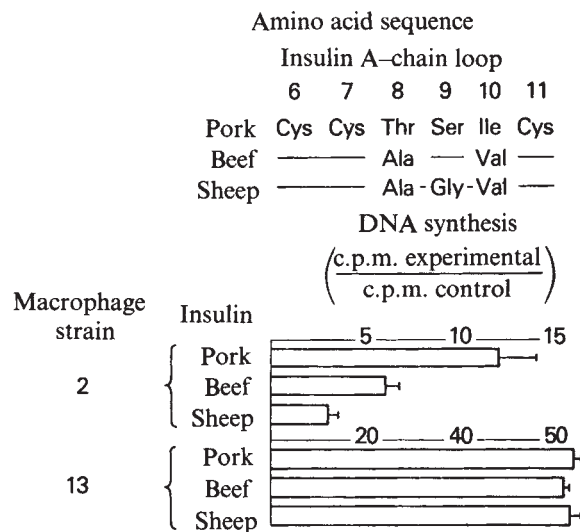


Fig. 1 Different response patterns observed in F₁ (2×13) guinea pig pork immune T cells when the antigen are 'presented' by strain 2 or strain 13 macrophages. In an experimental protocol similar to that described in Table 1 pork insulin immune F₁ (2×13) T cells were challenged 'in vitro' by strain 2 and strain 13 Mφ pulsed with pork, beef and sheep insulins. When the different insulins are presented by strain 2 Mφ, there was a good proliferative response to pork insulin, an intermediate response to beef insulin and almost no response to sheep insulin. There was a substantial proliferative response to the three different insulins when presented on strain 13 Mφ (b). Beef and pork insulins share one identity in the insulin A-chain loop and pork and sheep share no amino acids identities in this area (a). The amino acid sequence of the remainder of the beef, pork and sheep insulin molecule is identical.

strain 2 Mφ. This pattern of cross-reactivity is characterised by partial cross-reactivity between pork and beef insulin (partial identity in the loop) and little cross-reactivity between pork and sheep insulin (no identities in the loop). To corroborate the above findings and to show that a specific B-chain determinant is being selected when the whole insulin molecule is presented by strain 13 Mφ, we studied the response of oxidised insulin B chain immune F₁ (2×13) PELs to parental Mφ pulsed with native insulin or isolated B chain. Only strain 13 Mφ can present native insulin to B chain immune F₁ (2×13) PELs (Table 2). Despite the capacity of strain 2 Mφ to present the same insulin to insulin

Table 2 DNA synthetic response of T cells from oxidised B chain immunised F₁ (2×13) guinea pigs to parental macrophages pulsed with native insulin and isolated B chain

Macrophages pulsed with:	Macrophage genetic background	
	Strain 2	Strain 13
	³ H-TdR incorporation Δc.p.m. × 10 ⁻³	
PPD	145.81 ± 11.32	98.01 ± 14.10
Pork insulin	0.10 ± 0.09	13.16 ± 3.09
Oxidised B chain	6.84 ± 3.74	66.90 ± 19.60

Experimental procedures were as for Table 1 except F₁ (2×13) guinea pigs were immunised with 10 μg of oxidised insulin B chain 2–3 weeks before use. Data are mean ± s.e. for five experiments.

immune F₁ T cells, they cannot initiate DNA synthesis in B-chain immune F₁ T cells. Parenthetically, a considerable proliferative response is also seen when isolated insulin B chain is presented to F₁ T cells on strain 13 Mφ but less response is seen when the same antigen is bound to strain 2 Mφ.

Such studies raise the possibility that in F₁ (2×13) animals immunised to insulin, two distinct sets of clones of T cells are generated: one that recognises the α loop antigenic determinant 'processed' by or 'presented' by the strain 2 specific cell structures in the F₁ Mφ, and one that recognises

the B-chain determinant generated by the strain 13 counterpart of the system. This hypothesis has been examined by bromodeoxyuridine (BUdR) and light elimination experiments⁸ in F₁ guinea pigs.

Initial culture of insulin bearing parental macrophages with insulin immune F₁ T cells in the presence of BUdR and light selectively decreases responsiveness, on subsequent exposure to insulin bearing macrophages identical to those originally used to elicit cell activation (Table 3). Thus, preculture of F₁ T cells with strain 2 macrophages bearing insulin significantly depresses (65%) their response to antigen on strain 2 macrophages on subsequent cultures without comparably affecting responsiveness to insulin presented by strain 13 macrophages. Conversely, exposure of F₁ T cells to strain 13 macrophages bearing pork insulin eliminates 87% of responsiveness on repeat culture to pork insulin on strain 13 macrophages but does not affect the response of F₁ T cells to strain 2 macrophages bearing pork insulin. The clone remaining after elimination of T-cell proliferation by strain 2 macrophages pulsed with pork insulin responds equally well to strain 13 Mφ pulsed with either sheep (32,010 c.p.m.) or pork insulin (36,530 c.p.m.). Sheep insulin is identical to pork insulin except for the three amino acids of the A-chain α loop. Correspondingly, after BUdR and light elimination by strain 13 macrophages bearing insulin, the remaining cells respond only to strain 2 Mφ bearing pork but not sheep insulin.

Although immune specificity is traditionally considered a lymphocyte function⁸, previous observations from ours and from other laboratories indicate that the macrophage has some discriminatory function in the antigen recognition process (refs 3, 7, 9, 10, 12, 14, 15, 17).

Table 3 BUdR and light elimination of antigen specific F₁ T-cell proliferation: intramolecular determinant selection by pork insulin-bearing parental macrophages

Macrophages used in 1st culture	% Elimination of F ₁ T-cell DNA synthetic response on 2nd culture	
	Macrophage used:	
	Strain 2	Strain 13
Strain 2	65 (25-96)*	7 (0-12)
Strain 13	0	87 (70-97)

Immune PELs at 1×10^6 ml⁻¹ were cultured with 3×10^5 mitomycin C treated insulin pulsed Mφ for 48 h in 12 × 75-mm capped plastic tubes. At the end of this time 2μ g ml⁻¹ of freshly prepared BUdR was added to the cultures. Twenty-four hours later, the cell pellets were exposed to light by placing the culture tubes directly on an array of three fluorescent light bulbs (Cool-Ray GE) for 90 min. After four washes of the cell pellet with Hanks BSS, the BUdR and light-treated cells were cultured a second time in microtitre plates (as described in Table 1) in the presence of added pork insulin-pulsed Mφ. Data are expressed as % inhibition of DNA synthesis remaining after BUdR and light treatment. BUdR is a thymidine analogue which, if present during the S phase of the cell cycle, is incorporated into newly synthesised DNA and cross-links DNA strands upon light activation. Treatment of activated lymphocytes with BUdR and light leads to an irreversible block in cell replication.

*Values are averages of three experiments, with range in parentheses.

Although not excluding the possibility that Ir genes also function at the T-cell level, our data suggest that macrophages have a fundamental role in selecting, in a complex antigen, the moiety(ies) to be recognised by immune T cells. Indeed these experiments are of particular relevance in view of recent studies using BUdR and light to eliminate antigen specific T cells in populations of F₁ (2 × 13) lymphoid cells from guinea pigs immune to multi-determinant antigens¹⁸. Such experiments showed that elimination of T cell responsive to antigen associated with macrophages of one parental haplotype failed to eliminate responsiveness to the same antigen on the other parental macrophage. Those data were interpreted to indicate either: (1) that cells possess dual receptors, one for antigen and a second independent receptor for H-linked surface molecules or (2) that a

'compound antigen determinant' consisting of antigen in unique association with an Ia antigen not antigen alone, is recognised by a single T-cell receptor¹⁹. An additional interpretation raised by our studies of the immune response to insulin, is that a selected amino acid sequence and/or conformation within the antigen itself is seen by the T-cell receptor. Generation or display of such antigenic determinants would thus be a function of immune response genes operating at the level of the antigen presenting cell. Two general mechanisms by which macrophages might subserve such a function may be proposed. One possibility is that immune response genes define a class of receptors or broad specificity which recognise molecular shape and thus have the unique ability to focus or orient distinct regions of the antigen for presentation to the T cell. A second possibility is that immune response gene products are or regulate the activity of families of enzymes which modify or metabolise polypeptide antigens. In this latter situation, the repertoire of Ir genes associated with a given haplotype would define restricted areas of the molecule as available for display to the T-cell receptor. Experiments to distinguish these possibilities are in progress.

ALAN S. ROSENTHAL
MARCELLO A. BARCINSKI
J. THOMAS BLAKE

*Biologic Structure Section,
Laboratory of Clinical Investigation,
National Institute of Allergy and Infectious Diseases,
Bethesda, Maryland 20014*

Received 6 December 1976; accepted 10 March 1977.

- Shreffler, D. C. & David, C. S. *Adv. Immun.* **20**, 125-195 (1975).
- Erb, P., Meier, B. & Feldmann, M. *Nature* **263**, 601-604 (1976).
- Shevach, E. M. & Rosenthal, A. S. *J. exp. Med.* **138**, 1213-1229 (1973).
- Barcinski, M. A. & Rosenthal, A. S. *J. exp. Med.* **145**, 726-742 (1977).
- Waldron, J. A., Horn, R. G. & Rosenthal, A. S. *J. Immun.* **111**, 58-64 (1973).
- Rosenreich, D. L. & Rosenthal, A. S. *J. Immun.* **110**, 934-942 (1973).
- Rosenthal, A. S. & Shevach, E. M. *J. exp. Med.* **138**, 1194-1212 (1973).
- Zoschke, D. C. & Bach, F. H. *Science* **170**, 1404-1406 (1970).
- Benacerraf, B. & McDevitt, H. O. *Science* **175**, 273-279 (1972).
- Katz, D. H. & Benacerraf, B. *Transplant. Rev.* **22**, 175-195 (1976).
- Shevach, E. M. *J. Immun.* **116**, 1482-1489 (1976).
- Schwartz, R. H., David, C. S., Sachs, D. H. & Paul, W. E. *J. Immun.* **117**, 531-540 (1976).
- Greineder, D. K., Shevach, E. M. & Rosenthal, A. S. *J. Immun.* **117**, 1261-1266 (1976).
- Pierce, C. W., Kapp, J. A. & Benacerraf, B. *J. exp. Med.* **144**, 371-381 (1976).
- Miller, J. F. A. P., Vadas, M. A., Whitelaw, A. & Gamble, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2486-2490 (1976).
- Kappler, J. N. & Marrack, P. C. *Nature* **262**, 797-799 (1976).
- Thomas, D. W. & Shevach, E. M. *J. exp. Med.* **145**, 907-915 (1977).
- Paul, W. E., Shevach, E. M., Thomas, D. W., Pickeral, S. F. & Rosenthal, A. S. *Cold Spring Harb. Symp. quant. Biol.* **51**, 145, 618-638 (1977).
- Doherty, P. C., Blanden, R. V. & Zinkernagel, P. M. *Transplant. Rev.* **29**, 89-124 (1976).

Nucleoside-specific tolerance suppresses anti-nucleoside antibody forming cells

SINCE its original development by Jerne, the haemolytic plaque assay has increased our understanding of antibody formation to a wide variety of antigens, including proteins, lipopolysaccharides, and simple haptens². We have now developed an assay to detect plaque forming cells (PFC) making anti-nucleoside antibodies. Previously we reported the suppression of circulating antibody to DNA determinants by nucleoside-IgG conjugates^{3,4}. Here we show that BALB/c mice can be rendered tolerant in terms of both direct and indirect anti-nucleoside antibody forming cells and that the state of tolerance is nucleoside-specific at the cellular level.

Nucleosides were conjugated to BALB/c IgG2a (RPC₂) myeloma protein as tolerogen or to keyhole limpet haemocyanin (KLH) as immunogen by the procedure of Erlanger and Beiser⁵. Tolerogen preparations were (AGCT)₂₅-IgG2a and G₂₅-IgG2a and the immunogen preparation was (AGCT)₁₂₀-KLH; (the subscript numbers indicate the total molar ratio of hapten substitution on the

carrier, assuming 800,000 as the molecular weight of KLH).

To determine whether BALB/c mice immunised with 0.2 mg of AGCT-KLH in complete Freund's adjuvant intraperitoneally would produce antibody forming cells to AGCT-target sheep cells, groups of five mice were killed and assayed for direct and indirect PFC on various days after immunisation. The peak of direct anti-nucleoside PFC development occurred on days 6 and 7, and that of indirect

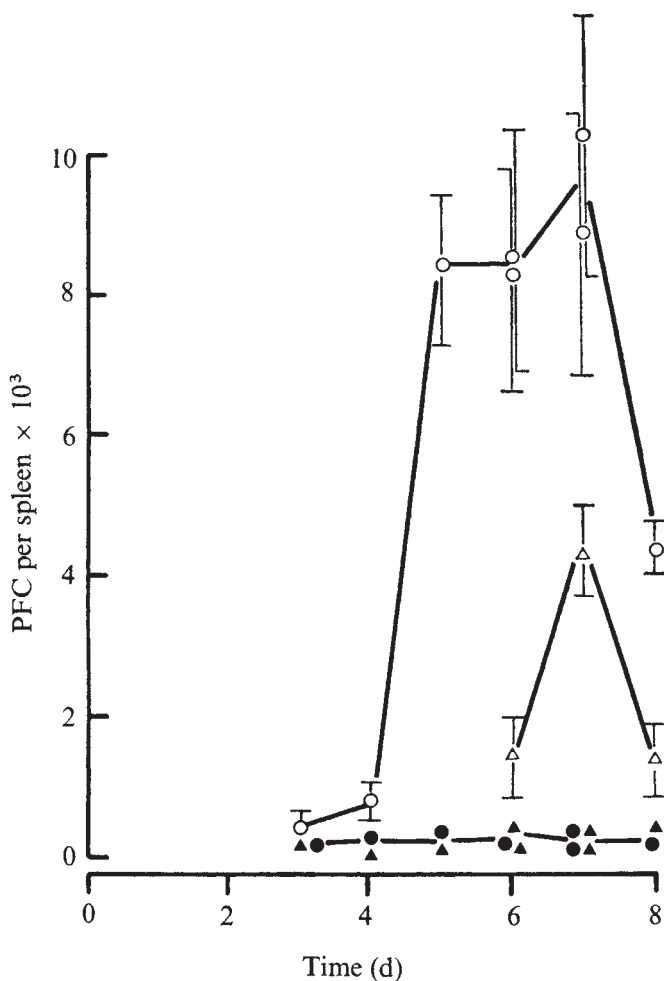


Fig. 1 Development of specific anti-nucleoside PFC, measured as the geometric mean ($\pm$ s.e.) PFC per spleen for groups of three non-immunised mice or five mice immunised with 0.2 mg of AGCT-KLH on day 0. Target cells were AGCT-SRBC. $\circ$, Direct and $\triangle$, indirect anti-nucleoside PFC in immunised mice. $\bullet$, Direct anti-nucleoside plaques in non-immunised mice. $\blacktriangle$, PFC to unsubstituted SRBC in immunised and non-immunised mice. The animals were 8–10 wk old BALB/c males (Jackson Laboratory). To prepare nucleoside-SRBC target cells, nucleosides were conjugated to the cells by the method of Erlanger and Beiser⁶ as modified to eliminate conditions that were damaging to the cells. A mixture of 5 mg of each nucleoside was oxidised in 1.5 ml of 0.1 M sodium periodate in 0.15 M NaHCO₃ for 20 min at room temperature; the reaction was stopped by the addition of 15 μ l of ethylene glycol. Sheep erythrocytes were washed twice with 0.15 M NaHCO₃ and 0.5 ml of packed cells were then suspended in 2.0 ml of the bicarbonate solution in a 40 ml centrifuge tube. The oxidised nucleosides were added dropwise to the cell suspension and the mixture was kept at room temperature for 15 min. 100 mg of *t*-butylamine borane (Aldrich) in 5 ml of 0.15 M NaHCO₃ was added. The suspension was kept at room temperature for 3 min and the tube was quickly filled with bicarbonate solution and centrifuged at 1,500 r.p.m. for 10 min. The modified cells were washed twice more with the bicarbonate solution and were ready for use as target cells in the haemolytic plaque assay of Jerne performed as before⁶. The indirect plaque forming cells were revealed by use of a rabbit anti-mouse IgG serum shown to react with IgG1, IgG2a, IgG2b, and IgG3 on immunodiffusion. The serum was used at a dilution of 1/200. IgG (7S) PFC were obtained by subtracting the direct PFC from the total number of PFC. Statistical analysis with Student's *t* test.

PFC on day 7 (Fig. 1). The spleens of non-immunised mice did not contain PFC to AGCT-SRBC spontaneously, and mice immunised with AGCT-KLH did not have a significant increase in the number of PFC to unsubstituted sheep red blood cells (SRBC).

To determine whether a nucleoside-IgG2a conjugate would suppress the appearance of anti-nucleoside antibody forming cells, mice were treated with 0.2 mg of tolerogen (AGCT-IgG2a) immediately before immunisation. They developed virtually no direct or indirect PFC (Fig. 2). Similar results were obtained when the tolerogen was given intravenously (i.v.) 10 d before immunisation (Table 1).

We then determined whether one can detect direct PFC to individual nucleosides in mice immunised with all four nucleosides, to allow a study of the nucleoside specificity of the tolerance. Seven BALB/c mice were injected with 0.2 mg of G-IgG2a. Ten days later, together with untreated controls, they were immunised intraperitoneally (i.p.) with 0.2 mg of AGCT-KLH in complete Freund's adjuvant, and the direct haemolytic plaque assay for all four nucleosides together (AGCT-SRBC), as well as for individual nucleoside SRBC targets, was performed 6 d later. Mice immunised with AGCT-KLH developed PFC to each of the four nucleosides (Table 2). The number of PFC was greatest for G and was much smaller for A than for any of the other nucleosides. Mice made tolerant by G-IgG2a had a specific suppression of antibody forming cells to this nucleoside only.

Several points are worth emphasis. First, when tolerance was induced in BALB/c mice by injection of nucleoside-IgG2a conjugates, the mechanism involved a reduction in the number of antibody forming cells and not simply the removal of preformed circulating antibody, suggesting that a similar mechanism may have occurred in (NZB/NZW)F1 mice made tolerant by nucleoside-IgG conjugates⁷. Second, when mice were given the G-IgG2a tolerogen and immunised with AGCT-KLH, they developed PFC to any nucleoside not present on the tolerance inducing carrier, demonstrating the exquisite specificity of tolerance at the cellular level. Third, tolerance was induced by nucleoside-IgG in terms of both direct and indirect PFC. It was also notable that while antibody forming cells to individual nucleosides could be detected readily in control mice

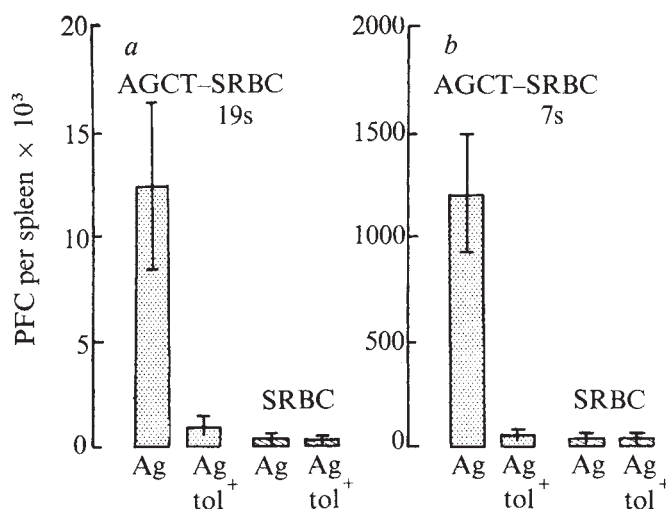


Fig. 2 Suppression of anti-nucleoside PFC formation by nucleoside-IgG conjugate. Each bar represents the geometric mean ($\pm$ s.e.) PFC per spleen of a group of five BALB/c mice either immunised i.p. with 0.2 mg of AGCT-KLH in complete Freund's adjuvant without pretreatment, or treated with 0.2 mg of AGCT-IgG2a immediately before the immunisation. Both AGCT-SRBC, prepared as described in the legend to Fig. 1, and unsubstituted SRBC were used as target cells in the haemolytic plaque assay done 6 d after immunisation. *a*, Direct PFC on day 6; *b*, indirect PFC on day 7.

Table 1 Effect of nucleoside-IgG2a conjugate on appearance of anti-nucleoside antibody forming cells

Treatment 10 d before immunisation	Direct (AGCT) PFC per spleen on day 7	<i>P</i>	Indirect (AGCT) PFC per spleen on day 7	<i>P</i>
None	10,173 ± 1,867		1,862 ± 594	
AGCT-IgG2a	1,888 ± 480	<0.001	498 ± 202	<0.01

Seven BALB/c mice were treated with 0.2 mg of AGCT-IgG2a, given i.v. and, together with seven untreated mice, were immunised i.p. with 0.2 mg of AGCT-KLH in complete Freund's adjuvant 10 d later. Direct and indirect PFC were measured on day 7, as described in the legend to Fig. 1. Values are means ± s.e.

Table 2 Nucleoside specificity of tolerance

Treatment 10 d before immunisation	AGCT-SRBC	<i>P</i>	Direct PFC per spleen on day 6		<i>P</i>	C-SRBC	T-SRBC
			A-SRBC	G-SRBC			
None	19,970 (7,283)		1,055 ± 922	8,848 ± 3,538		6,706 ± 3,730	7,943 ± 4,546
G-IgG2a	12,751 (7,283)	>0.4	1,535 ± 689	883 ± 710	<0.01	5,824 ± 4,791	6,067 ± 2,066

Seven BALB/c mice were treated with 0.2 mg of G-IgG2a and, together with seven untreated mice, were immunised i.p. with 0.2 mg of AGCT-KLH in complete Freund's adjuvant 10 d later. Individual nucleoside-SRBC target cells were prepared as in the legend to Fig. 1. Values are means ± s.e.

immunised with all four nucleosides on the immunogenic carrier, there was some variation in the number of PFC specific for each nucleoside. This reflected at the cellular level of the immunodominance of guanosine and the low immunogenicity of adenosine that we have noted previously for serum antibody responses⁴. Finally, from a technical point of view, the modifications of the technique of Erlanger and Beiser⁵ to link nucleosides to sheep cells should be noted. The ionic strength had to be maintained to prevent lysis. The use of 0.15 M bicarbonate did this and provided a suitable pH for the Schiff base formation. The use of *t*-butylamine borane complex in place of sodium borohydride was important in preventing damage to the erythrocytes during reduction of the Schiff base, and the reduction time was brief.

The immune response to nucleic acid antigens is important from at least two points of view. First, since anti-nucleic acid antibodies are involved in autoimmune disease⁶, the suppression of this response can be of therapeutic importance; and second, strain differences in the ability of mice to respond to nucleic acid antigens have been observed⁹ and might offer the opportunity to study the genetic basis of both immunity and tolerance to defined determinants of naturally occurring antigens. The development of an assay that permits the detection of antibody forming cells to either all four nucleosides simultaneously or to individual nucleosides is useful for the study of these phenomena.

This work was supported by the US NSF, NIH and the Massachusetts Lupus foundation. We thank Ms L. Kilham for technical assistance.

B. DAVID STOLLAR

Department of Biochemistry and
Pharmacology,
Tufts University School of Medicine,
Boston, Massachusetts 02111

YVES BOREL

Division of Immunology,
The Children's Hospital Medical Center,
Department of Pediatrics,
Harvard Medical School,
Boston, Massachusetts 02115

Received 2 February; accepted 10 March 1977.

¹ Jerne, N. K., Nordin, A. A. & Henry, C. L. in *Cell Bound Antibodies* 109-116 (Wistar Institute, Philadelphia, 1963).

² Cunningham, A. J. *Progr. Allergy* 17, 5-50 (1973).

³ Stollar, B. D. & Borel, Y. J. *Immun.* 115, 1095-1100 (1975).

⁴ Stollar, B. D. & Borel, Y. J. *Immun.* 117, 1308-1313 (1976).

⁵ Erlanger, B. F. & Beiser, S. M. *Proc. natn. Acad. Sci. U.S.A.* 52, 68-74 (1967).

⁶ Golan, D. T. & Borel, Y. J. *exp. Med.* 134, 1046-1061 (1971).

⁷ Borel, Y., Lewis, R. M. & Stollar, B. D. *Science* 182, 76-78 (1973).

⁸ Koffler, D. J. *exp. Med.* 126, 607-623 (1967).

⁹ Fuchs, S., Mozes, E. & Stollar, B. D. *J. Immun.* 114, 1287-1291 (1973).

T-cell cytotoxicity in the absence of viral protein synthesis in target cells

CYTOTOXIC T cells lyse only those virus infected target cells *in vitro* which express, in addition to the viral antigen(s), those *K* or *D* region products of the major histocompatibility complex (MHC) which were present during anti-viral sensitisation *in vivo*. This 'associative recognition' by cytotoxic T cells could reflect the interaction of two T-cell receptors with specificity for target *K* or *D* gene products and independently for the viral antigen, or one receptor

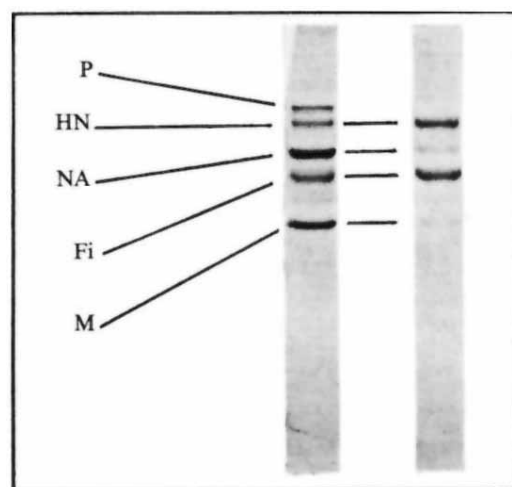


Fig. 1 Separation of Sendai virus envelope protein on SDS-polyacrylamide gels. Sendai virus (8 mg ml⁻¹) in phosphate-buffered saline without Mg or Ca (PBS-A), 1.5 × 10⁵ haemagglutinating units (HAU) per ml were treated with 0.25% NP 40 for 15 min at 20 °C (ref. 6). The suspension was centrifuged at 120,000g for 1 h at 4 °C to remove partially disrupted virions and nucleocapsids. The supernatant was dialysed for 4 d at 4 °C against twice daily changes of PBS-A using spectrophor 2 tubing (Raven Scientific). The reassembled membrane particles were concentrated by centrifugation at 120,000g for 1 h at 4 °C and the pellet was resuspended in PBS-A. Envelopes were active in assays for haemagglutination⁷, neuraminidase⁸ and haemolytic activities⁹. Envelopes were analysed by SDS-polyacrylamide electrophoresis by the method of Laemmli¹⁰ using 10% gels. P, polymerase; HN, haemagglutinin; NA, Neuraminidase; F₁, fusion factor; M, M-protein.

with specificity for virally altered *K* or *D* region products (see ref. 1 and refs therein). There are various ways that the MHC antigens could be altered, including 'modification from within', where the virus modifies host protein synthesis by interfering with transcription², translation or post-translational glycosylation; or 'modification from without' where enzymic or chemical alteration of cell membrane proteins are induced by virus activity at the cell surface. In this report we show that inactivated Sendai virus or isolated Sendai virus envelopes can serve to modify a cell and make it a specific target for Sendai-immune T-cell killing, thus excluding the possibility of 'modification from within' in this system.

Sendai virus-primed BALB/c spleen cells, boosted *in vitro* with syngeneic infected blast cells, specifically killed Sendai virus-infected syngeneic target cells; primed spleen cells boosted with syngeneic non-infected cells did not (Table 1). Killer cells after secondary stimulation were T cells as shown by treatment with anti-Ly-sera (U. K. and P. Beverley, unpublished). Three approaches were used to show that synthesis of viral protein is not required in the target cell in order for it to be killed by virally sensitised T cells. First, cells incubated with ultraviolet-irradiated Sendai virus were lysed by sensitised T cells (data not shown). Second, cells exposed to purified Sendai virus, inactivated with β -propiolactone—which abolishes induction of protein synthesis by inactivated particles⁵—also proved to be effective target cells (Table 1). Treatment with β -propiolactone reduced the number of plaque forming units by 10^6 . Since the ratio of particles to plaque forming units would be expected to be no better than $10:1$ (J. Skehel, personal communication) and target-to-particle ratio was $1:10^6$, it follows that, at most, one in a thousand cells were infected with a live virus. Sendai virus specific killer cells generated by secondary stimulation *in vitro* lyse only *H-2*-compatible infected target cells (U. K., unpublished) and the same restriction is seen using target cells incubated with inactivated Sendai virus (Table 2). Third, cells exposed to isolated envelopes from Sendai virus, which contained mainly the two surface glycoproteins with a very small amount of nucleocapsid and almost no detectable M-protein (Fig. 1), were susceptible to T cell-mediated lysis. To exclude nonspecific effects of the Sendai envelopes on target cells, influenza virus-specific cytotoxic T cells were tested on these target cells, and found to be ineffective (Table 3).

In several virus infections, target cell formation requires translation of early viral proteins¹¹⁻¹³, glycosylation¹⁴, and

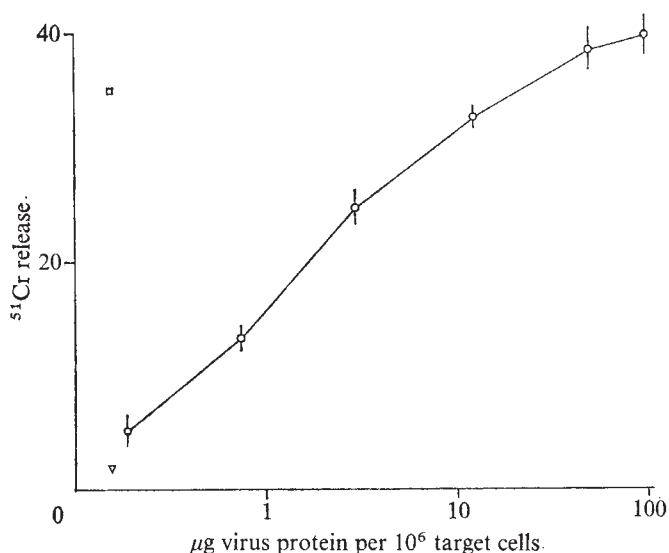


Fig. 2 T cell-mediated lysis of P-815 target cells exposed to dilutions of purified Sendai virus inactivated with β -propiolactone (B-PL). Sendai virus was grown in the allantoic cavity of 10-d-old embryonated fowl eggs and purified by centrifugation through a sucrose gradient (20–55% in PBS-A). The virus-containing band was collected and concentrated by centrifugation at 120,000g for 1 h. Inactivation was achieved by incubating the purified virus (2 mg per ml, 4×10^4 HAU per ml, $1 \mu\text{g} \approx 10^{11}$ virus particles) for 18 h at 4 °C with β -propiolactone (Sigma) at a concentration of 0.025%. β -propiolactone was inactivated by incubation for 30 min at 37 °C. Inactivation was tested by egg infectivity titration¹⁵. ○—○, Lysis of P-815 cells exposed 2 h before test with dilutions of B-PL-Sendai virus (100 μg virus per 10^6 cells, about 10^5 particles per cell). □, P-815 targets infected overnight with 20 PFU per cell (1 PFU ≤ 10 particles). ▽, Non-infected P-815 control cells. Assay conditions identical to Table 1, incubation time 3 h, ratio of killer-target cells, 5:1.

expression of viral proteins on the cell surface. The ability of Sendai virus to by-pass these processes may be due to its unique capacity to mediate cell-cell fusion, which could allow the insertion of large amounts of virus glycoprotein into the target cell membrane. From the dose-response curve (Fig. 2) of target cell formation with dilutions of β -propiolactone inactivated Sendai virus, it can be seen that about 10^6 particles per target cell were required to induce

Table 1 Lysis of syngeneic target cells incubated with infective and inactivated Sendai virus or vesicles containing viral envelope proteins

Responder cells	Stimulator cells	Culture day	Ratio K:T	Targets			
				P-815	P-815/Sendai	P-815/B-PL-S	P-815/Sendai env.
BALB/c, Sendai immune	BALB/c con A treated blasts, Sendai infected	5	10:1	1.1	43.5	42.0	18.1
			5:1	0.3	26.6	26.3	9.7
			2.5:1	1.7	13.9	15.1	6.3
BALB/c, Sendai immune	BALB/c con A treated blasts	5	10:1	2.8	0.5	1.3	0.4
			5:1	3.5	0.9	0.9	1.0
BALB/c, Normal	BALB/c con A treated blasts, Sendai infected	5	10:1	0.1	0.9	4.0	1.6
			5:1	0.6	0.4	3.6	0.5

BALB/c mice were immunised to Sendai virus by i.p. injection of 0.2 ml allantoic fluid containing 10^7 plaque forming units (PFU) of virus⁹. Six weeks later spleens were collected and incubated in RPMI medium containing 10% foetal calf serum, glutamine and 5×10^{-5} M 2-mercaptoethanol in 2 ml well linbro plates in presence of stimulated cells. Stimulator cells were prepared by incubation of spleen cells in culture medium containing $2.5 \mu\text{g ml}^{-1}$ concanavalin A (con A) for 48 h followed by overnight infection with 20 PFU per cell Sendai virus and 2,000 rad irradiation. 5×10^6 responder cells were incubated with 5×10^6 stimulator cells. Cytotoxic T cells were collected from cultures 4–7 d later⁴ and a ⁵¹Cr release assay was performed with 5×10^3 syngeneic tumour target cells per well⁸. Target cells were a, ⁵¹Cr labelled non-infected P-815 cells; b, P-815 cells infected overnight with 20 PFU per cell Sendai virus; c, P-815 cells pre-incubated for 2 h with β -propiolactone (B-PL) inactivated (see legend Fig. 2), purified Sendai virus in a protein concentration of 50 μg per 10^6 target cells; d, P-815 cells pre-incubated for 2 h with Sendai envelope (env.) proteins (see legend to Fig. 1) in a protein concentration of 50 μg per 10^6 target cells. Results are from triplicate assays; S.E.M. were always less than 3. Spontaneous lysis, which did not exceed 15–20%, is subtracted.

Table 2 Lysis of target cells treated with β -propiolactone inactivated Sendai virus is *H*-2 restricted

Responder cells	Stimulator cells	P-815 S-B-PL (<i>H</i> -2 ^{dd})	L-929 S-B-PL (<i>H</i> -2 ^{kk})	EL-4 S-B-PL (<i>H</i> -2 ^{bb})
BALB/c normal (<i>H</i> -2 ^{dd})	BALB/c con A treated blast, Sendai infected	3.1	1.4	2.5
BALB/c immune (<i>H</i> -2 ^{dd})	BALB/c con A treated blasts, Sendai infected	46.1	1.5	11.9
CBA immune (<i>H</i> -2 ^{kk})	CBA con A treated blasts, Sendai infected	2.4	30.6	3.0
C57BL6 immune (<i>H</i> -2 ^{bb})	C57 BL6 con A treated blast, Sendai infected	—	3.8	32.5

Conditions of immunisation and *in vitro* restimulation were as for Table 1. Cells were collected and tested on culture day 6. Target cells were incubated with B-PL-virus, in a protein concentration of 50 μ g per 10⁶ target cells. Results are from triplicate assays, ratio of killer–target cells is 20:1, incubation time 4 h.

a comparable degree of killing as seen with infected target cells. Since the percentage of particles which actually fuse with the target cells is unknown, the number of fused particles per cell necessary for target cell formation cannot be estimated. Failure to use sufficiently high concentrations of virus during target cell incubation may explain previous negative findings (ref. 12, and U.K., unpublished).

β -propiolactone-Sendai virus-treated target cells were lysed to a similar extent as target cells incubated with live virus, but target cells which were produced with Sendai virus envelope vesicles were consistently lysed at a lower plateau level. Two explanations are possible. There could be two populations of Sendai virus specific cytotoxic T cells

within'. If there are alterations of *H*-2 gene products these must occur from without, at the cell surface. Envelopes contain neuraminidase activity, but cleavage of neuraminic acid seems not to change the *H*-2 surface antigen specificity^{16,17}. We do not know if there are other residual enzyme activities such as proteases or glycosidases remaining in the envelopes. It is also possible that cross-linking of sialic acids by the viral haemagglutinin might change the surface distribution of *H*-2, and sometimes create effective target antigens.

Which then are the viral antigens on cell surfaces recognised by the T cells? Recent studies on influenza virus-specific CTL have suggested that the M-protein which is

Table 3 Virus-specific lysis of target cells incubated with inactivated Sendai virus or Sendai virus envelopes

Responder cells	Stimulator cells	Culture day	Ratio K:T	Targets				
				P-815	P-815/influenza	P815/Sendai	P-815/B-PL-S	P-815/Sendai env.
BALB/c, Sendai immune	P-815, Sendai infected	5	10:1	4.3	2.3	39.6	28.1	19.3
			5:1	–2.5	3.0	37.8	19.8	14.2
BALB/c, influenza immune	P-815, influenza infected	5	10:1	–0.9	27.4	0.8	1.3	1.8
			5:1	0.2	13.9	0.3	–3.8	3.9
(C3H $\times$ DBA/2) F ₁ , Sendai immune	BALB/c LPS blasts, Sendai infected	5	10:1	–0.29	3.7	26.5	27.2	—
			5:1	1.02	1.4	8.6	8.5	—
(C3H $\times$ DBA/2) F ₁ , influenza immune	BALB/c LPS blasts, influenza infected	5	10:1	0.3	19.8	3.4	3.8	—
			5:1	2.9	11.3	0.5	2.9	—

Responder cells (BALB/c, (C3H $\times$ DBA/2) F₁) were collected from mice injected i.p. 6 weeks previously with 10⁷ PFU Sendai virus or influenza virus. The influenza strain A/Jap/Bel/305/57 (H₂N₂) a recombinant of A/Jap/305/57 (H₂N₂) and A/Bel/42 (H₂N₂) (kindly supplied by Dr J. Skehel) was used. Conditions of *in vivo* sensitisation and *in vitro* boosting of cells were similar to those used for secondary cytotoxic T-cell generation against Sendai virus. Stimulator cells were BALB/c lymphocyte blasts after 48 h incubation in 50 μ g ml^{–1} lipopolysaccharide containing medium or P-815 tumour cells. Stimulator cells were infected overnight with 20 PFU per cell and irradiated with 2,000 rad (blasts) or 8,000 rad (P-815). Ratio of stimulator to responder cells and ⁵¹Cr release assay conditions were identical to those described in the legend to Table 1.

with receptors for different viral antigens; however, antigens on envelope-treated target cells are at least partially shared with those of infected target cells since the latter compete in cold cell inhibition tests with the lysis of target cells exposed to Sendai virus envelopes (data not shown). The second possibility is that the envelope vesicles are less active in fusing with the target cell membrane; this is the more likely explanation since the β -propiolactone inactivated virus was active in cell–cell fusion, while the envelopes were not.

These results rule out, at least for Sendai virus infection, formation of an 'altered self' by a 'modification from

not glycosylated and not expressed on cell membranes may contribute to target specificity¹⁸. This is very unlikely for Sendai virus since the envelope preparations contain only two glycoproteins, one with neuraminidase and haemagglutinin activity and the other with fusion activity.

These results show for the first time that target cells for virus-specific T-cell lysis can be prepared with vesicles containing essentially only viral surface glycoproteins and that viral protein synthesis is not required for target cell formation.

We thank Drs Martin Raff and Peter Beverley for helpful discussion and Rowan Coyle and John Sever for technical

assistance. U.K. was supported by Deutsche Forschungsgemeinschaft and M.W. by the American Heart Association.

U. KOSZINOWSKI

ICRF Tumour Immunology Unit,
Department of Zoology,
University College London,
Gower Street, London WC1, UK

M. J. GETHING
M. WATERFIELD

Imperial Cancer Research Fund Laboratories,
Lincoln's Inn Fields,
London WC2, UK

Received 1 February; accepted 14 March 1977.

- 1 Munro, A. & Bright, S. *Nature* **264**, 145-152 (1976).
- 2 Garrido, F., Schirmacher, V. & Festenstein, H. *Nature* **259**, 228-229 (1976).
- 3 Ertl, H. & Koszinowski, U. *Immunobiology* **152**, 128-140 (1976).
- 4 Gardner, I. D. & Blanden, R. V. *Cell Immun.* **22**, 283-296 (1976).
- 5 Neff, J. & Enders, J. F. *Proc. Soc. exp. Biol. Med.* **127**, 260-267 (1968).
- 6 Hosaka, Y. & Shimizu, Y. K. *Virology* **49**, 627-639 (1972).
- 7 Salk, J. J. *Immun.* **49**, 87-98 (1944).
- 8 Aminoff, D. *Biochem. J.* **81**, 384-392 (1961).
- 9 Shimizu, Y. K. *J. Virol.* **9**, 842-852 (1972).
- 10 Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- 11 Ueda, Y. & Tagaya, I. *J. exp. Med.* **138**, 1033-1043 (1973).
- 12 Ada, G. I., Jackson, D. C., Blanden, R. V., Tha Hia, R. & Bowern, N. A. *Scand. J. Immun.* **5**, 23-30 (1976).
- 13 Koszinowski, U. & Ertl, H. *Eur. J. Immun.* **6**, 679-683 (1976).
- 14 Jackson, D. C., Ada, G. I., Hapel, A. J. & Dunlop, M. B. C. *Scand. J. Immun.* (in the press).
- 15 Scheid, A. & Chopin, P. W. *Virology* **62**, 125-133 (1974).
- 16 Sandford, B. H. & Codrington, J. F. *Tissue Antigens* **1**, 153-161 (1971).
- 17 Ray, B. K. & Simmons, R. L. *Proc. Soc. exp. Biol. Med.* **142**, 846-852 (1973).
- 18 Effros, R. B., Doherty, P. C., Gerhard, W. & Bennink, J. J. *exp. Med.* (in the press).

Cellular inactivation by ultrasound

THE lethal effect of ultrasound (US) on mammalian cells has received relatively little attention. Understandably, potential genetic aspects of US have been of prime concern to physicians who use US as a diagnostic tool; at the average power densities involved ($\ll 1 \text{ W cm}^{-2}$) little, if any cell killing is to be expected. There have been sporadic attempts to use higher intensities ($\sim 1 \text{ W cm}^{-2}$) as a treatment modality in cancer therapy^{1,2}, but those experiments seem to have been based on inadequate cellular studies. The effects of US usually were evaluated in terms of morphological criteria rather than on quantitative determination of the loss of viability as measured by colony formation. There are few reports of the effects of US on survival of mammalian cells^{3,4}, and none specifically examine hyperthermic interaction. With the increased interest in hyperthermia for tumour therapy, attention has been directed towards the use of ultrasound to achieve tumour heating. In preliminary experiments in which US was used to heat the EMT6 sarcoma and KHJJ carcinoma in mice, we found a high percentage of tumour cures with short ($\sim 30 \text{ min}$) treatments at temperatures ($43\text{--}44^\circ\text{C}$) where *in vitro* results of hyperthermia-induced cell killing⁵⁻⁷ would not have led to a prediction of any cures. We therefore initiated an investigation of the effects of US on survival of Chinese hamster cells to see if direct cell killing by US could explain our *in vivo* results, or, as in the case of radiofrequency (RF) electromagnetic heating, we would be forced to invoke host response⁸. In particular, we examined the thermal and non-thermal components of cellular inactivation by US. We report here that there is a definite non-thermal cytotoxic effect of US. Its relative contribution to cell killing is a highly nonlinear function of the temperature of the cellular milieu. The survival curves show clearly that, beyond an initial threshold, small changes in temperature and/or US intensity can give rise to impressive changes in survival values. The threshold nature of the data strongly suggests that by means of overlapping beams, ultrasound energy could be delivered to tumour tissue to achieve massive cell killings while

sparing normal tissue outside the tumour volume to a degree far exceeding that of conventional techniques.

Chinese hamster cells (HA-1) of ovarian origin were maintained in Eagle's MEM supplemented with 15% foetal calf serum, penicillin, and streptomycin. The cultures were kept in a humidified incubator with a mixture of 95% air and 5% CO_2 , and were routinely checked for mycoplasma. To maximise the transmission of the acoustic energy and hence minimise heat induced by absorption of US, cells were grown as monolayers in specially designed stainless steel dishes with Mylar bottoms ($130 \mu\text{m}$ thick). Approximately 8×10^4 cells were seeded within a confined area (1.6 cm in diameter) on the thin Mylar film; experiments were always performed on the third day after plating when the cell density reached approximately $5 \times 10^5 \text{ cells cm}^{-2}$ and the cells were still in exponential growth. Plating efficiencies of cells trypsinised from the Mylar film were 70-90%. Some experiments were performed with cells in suspension; as far as possible, the geometry was identical to that used for attached cells. The cloning technique of Puck and Marcus was used to determine the fraction of viable cells. The US irradiating system consisted of a laboratory-developed signal generator and power amplifier driving a piezo-electric transducer (Mettler Electronics). The transducer was mounted on the frame

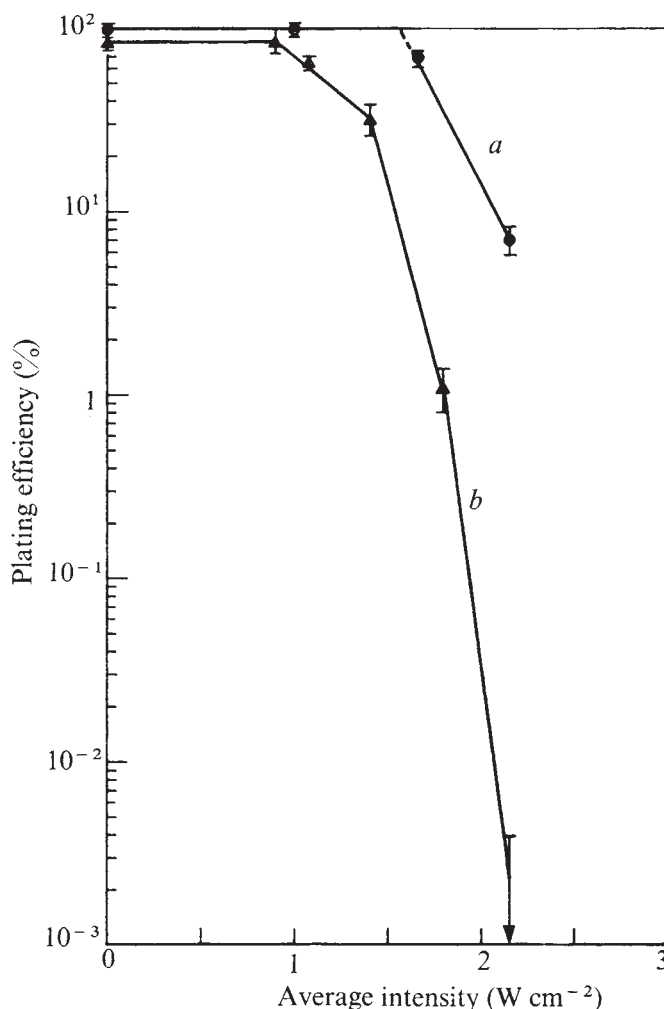


Fig. 1 Plating efficiencies (normalised to untreated controls) against US intensity. Cells were irradiated for 30 min at a, 37°C ; b, 43°C . Error bars in this and the succeeding figures show s.e.m. Plating efficiency is defined as (no. colonies per dish/no. cells plated per dish) $\times 100$. Cells plated per dish represents countable cells obtained by enzymic removal from Mylar after treatment.

of a specially designed dish holder, in a circulating hot water bath (Forma-Temp. Jr), so that alignment between cells and the transducer was accurately reproducible. The Mylar bottom of the dish was always in direct contact with circulating water of the desired temperature. The cells were irradiated at a distance of 13 cm from the transducer (3.5 cm diameter). The operating frequency was 1 MHz and the signal was pulse-modulated at about $1,000 \text{ cycles s}^{-1}$ (duty cycle 60%). The irradiation field was 1.6 cm in diameter at this source-target distance; the spatial uniformity of the field intensity was within $\pm 20\%$ as measured by an absorbing probe. This probe consisted of a 22 gauge needle with a thermistor mounted inside the top; the tip was surrounded by a cylinder of silicon rubber (3 mm in diameter and 5 mm in length). The probe was calibrated by RCA laboratories, Princeton, New Jersey; the intensity measured included any reflected contribution. Agitation minimised any standing wave formation, however. The temperature of the water bath was maintained to $\pm 0.1^\circ\text{C}$; the temperature difference between the circulating water and the medium inside the dish was less than 0.1°C even at US intensities of 2 W cm^{-2} . The time required to reach thermal equilibrium was about 2–3 min; this is included in the quoted times of heating and/or US irradiation. Temperature measurements in the presence of US fields were obtained using thermocouples encased in 28 gauge needles. At intensities of approximately 0.7 W cm^{-2} at the measuring instrument, we saw no interaction of US and the sensor. When higher intensities were used, such interaction was observed and had to be taken into account; an infrared pyrometer was then used to measure the temperature of the medium inside the dish.

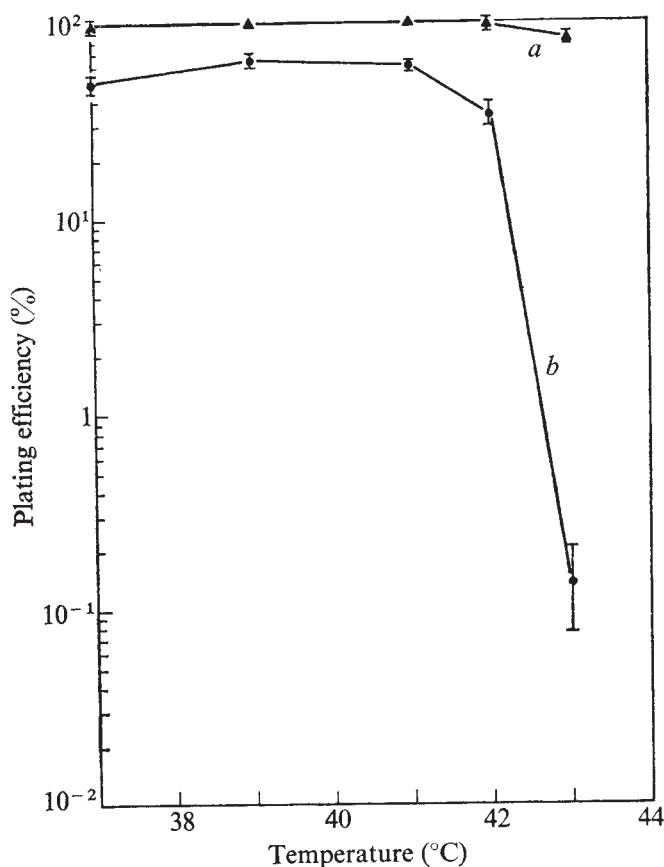


Fig. 2 Plating efficiencies (normalised to untreated controls) against temperature. *a*, Cells exposed for 30 min to heat alone; *b*, cells irradiated for 30 min with US (2 W cm^{-2}) at the indicated temperatures.

For each experiment, survival values of cells exposed to heat treatment in the absence of US irradiation were measured; the results were all in good agreement with previously published data⁶ and were used as thermal controls. Any cytotoxicity above these control values is attributed to non-thermal mechanisms.

Monolayers of cells were exposed to pulsed US radiation for 30 min at various average intensities and at water bath temperatures ranging from 37 to 45°C . Plating efficiencies (all normalised to those of untreated controls) of cells irradiated at 37°C and 43°C are shown in Fig. 1. The cytotoxicity of US irradiation had a threshold when plotted against average acoustic intensity. This threshold intensity was temperature dependent: at 43°C it was lower than

Table 1 Effect of ultrasound radiation on cell number

Average intensity (W cm^{-2})	Countable cells on Mylar after 30 min US irradiation (% of unirradiated controls)	
	37°C	43°C
0.75	100	100
1.5	18	20
2.0	5	5

37°C . Cells were then exposed at various temperatures either with or without pulsed ultrasound at an average intensity of 2 W cm^{-2} . Figure 2 shows the plating efficiencies as a function of water bath temperature. Survival of cells exposed simultaneously to heat and ultrasound was lower than that of cells exposed to heat alone. Plotted against temperature, the cytotoxicity of US irradiation at a constant average intensity again showed a threshold.

In addition to changes in plating efficiency, there was also a reduction in cell number, showing the occurrence of immediate cell lysis. Table 1 indicates the percentage of cells which were attached to the Mylar after 30 min exposure to US, with respect to the untreated control. The cell number includes both viable and non-viable cells. In contrast to plating efficiency, the fraction of cells lysed showed no significant temperature dependence between 37 and 43°C . Survival values, after correction for cell lysis, are shown in Fig. 3 for one set of experiments. The number of cells recovered from the supernatant was negligible.

The possibility that these results were artefact of the attachment of cells to the Mylar was ruled out by the finding that cells in suspension responded similarly to US irradiation (data not shown). The nature of US-induced cell damage is not clearly understood. In general, US effects are classified under thermal, cavitation or 'other'^{9–11}. In the experiments described above, the magnitude of thermal effects has been evaluated independently, since the macroscopic temperature in each case was established by the hot water bath and not by the US intensity. The cell killing over and above that caused by this temperature is clearly due to non-thermal effects. These non-thermal effects may be divided into temperature-dependent and temperature-independent components. Cell lysis is largely temperature-independent, while cell inactivation, as measured by the reduction in plating efficiency, results at least in part from temperature-dependent effects. Cavitation played an important part; this was shown in experiments in which the average US energy was kept constant, but pulse duration and repetition frequency were varied. With pulses of less than 1 ms, much less cell lysis and cell death were seen; 1 ms is about the minimum time required for the onset of cavitation¹². Even much shorter pulses induced considerable lysis and cell death, however; hence other phenomena such as, for example, acoustic streaming were involved.

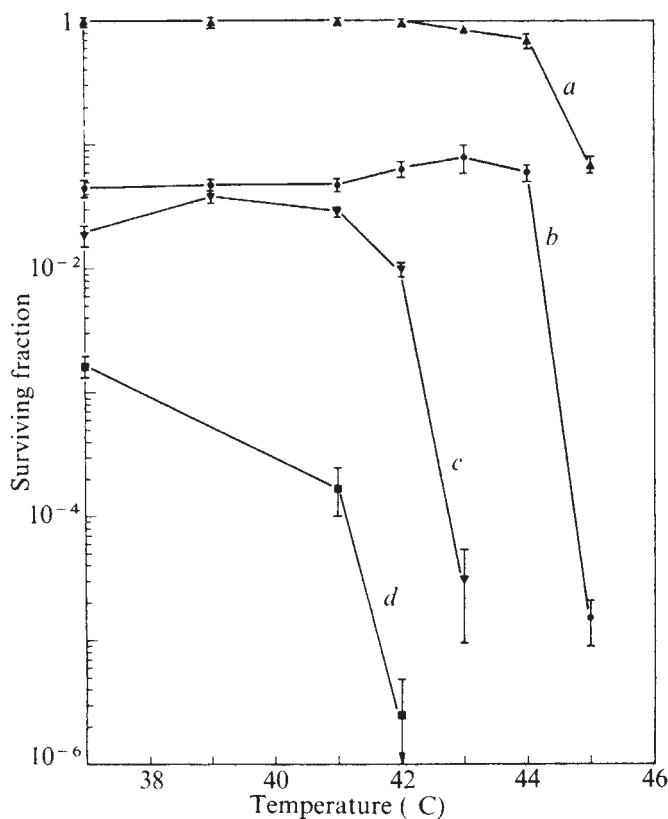


Fig. 3 Surviving fractions of cells exposed to US for 30 min at various temperatures and intensities. *a*, Heat alone; *b*, 1 W cm⁻²; *c*, 2 W cm⁻²; *d*, 2.2 W cm⁻². Surviving fraction (SF) is defined as (cell no. after treatment/cell no. before treatment) × (plating efficiency).

The dramatic temperature-dependent changes in plating efficiencies observed (Figs 1 and 2) are unlikely to be related to changes in the physical action of ultrasound, but rather to intracellular changes. Experiments in ours and other laboratories have demonstrated a variety of temperature-related changes in the response of cells to cytotoxic agents¹³, to hypo- or hypertonic environments¹⁴, and to membrane-specific antibiotic amphotericin B (ref. 14). These changes, which probably involve modification of cellular membranes, occur in the range of 41–43 °C, and all are in the direction of appreciably and sometimes dramatically greater sensitivity to the external challenge. Thus, it is not unreasonable to find that the cells' response to US also changes drastically with temperature. It must be kept in mind, however, that the observations of thermal interactions depend critically on the particular parameters of US irradiation. Intensity variations, frequency variations (or modulation) and changes in pulse width can shift the thresholds appreciably. This makes it imperative that, for comparison of results, conditions be standardised, and that the details of US irradiations be reported carefully.

The shapes of the survival curves (Figs 1–3) suggest that US may be a very useful tool in the eradication of localised tumours. Since US can readily be focused, it should not be difficult to obtain 'dose distribution' such that temperature and intensity thresholds are exceeded only within a hypothetical volume. Within that volume, extensive cell killing would then be effected, while in the surrounding normal tissue very little cytotoxicity would be expected. In that sense, then, US would differ greatly from X-ray (or even high LET particle) therapy, as well as from focal ultrasound ablation¹⁵, where normal tissue toxicity outside the tumour volume is frequently the treatment-limiting consideration. It would also differ from RF-induced current

heating, where cell inactivation seems to be largely (or exclusively) through thermal mechanisms⁸. The tumour experiments referred to earlier are consistent with this expectation: in mice which had been cured of their tumours by US and which were sacrificed as late as 180 d after treatment, little or no evidence of normal tissue damage was seen (Fajardo, Marmor and Hahn, unpublished). We are not sure, however, that the tumour cures can be explained entirely on the basis of this direct US-related cell killing: at the intensity levels of the tumour irradiations, at least the peripheries of the tumours should have been below the cytotoxic threshold levels found *in vitro*.

We thank Drs R. S. Mezrick and D. H. R. Vilkomerson for calibrating our probes, and Mr D. Pounds and Ms E. Shiu for technical assistance. This work was supported by US Public Health Service grants.

GLORIA C. LI
GEORGE M. HAHN

Department of Radiology,
Stanford University School of Medicine,
Stanford, California 94305

LEONARD J. TOLMACH

Department of Anatomy,
Washington University School of Medicine,
St Louis, Missouri 63130

Received 29 November 1976; accepted 8 March 1977.

- 1 Woebber, K. in *Ultrasonic Energy* (ed. Kelly Fry, E.) 137–149 (University of Illinois Press, Urbana, 1965).
- 2 Horvath, J. *Strahlentherapie* 75, 119–125 (1944).
- 3 Clarke, P. R. & Hill, C. R. *J. acoust. Soc. Am.* 47, 649–653 (1970).
- 4 Kaufman, G. E., Miller, M. W., Dan Griffiths, T., Ciaravino, V. & Carstensen, E. L. *Ultrasound Med. Biol.* (in the press).
- 5 Hahn, G. M. *Cancer Res.* 34, 3117–3123 (1974).
- 6 Westra, A. & Dewey, W. C. *Int. J. Radiat. Biol.* 19, 467–477 (1971).
- 7 Marmor, J. B., Hahn, G. M. & Pounds, D. *Cancer Res.* (in the press).
- 8 Marmor, J. B., Hahn, N. & Hahn, G. M. *Cancer Res.* 37, 879–883 (1977).
- 9 Hughes, D. E. & Nyborg, W. L. *Science* 138, 108–113 (1962).
- 10 Rooney, J. A. in *Interaction of Ultrasound and Biological Tissues* (eds Reid, J. M. & Sikov, M. R.) DHEW Publication (FDA) 73-8008 BRH/DBE 73-1, 93–97 (1972).
- 11 Nyborg, W. L. *J. acoust. Soc. Am.* 44, 1304–1309 (1968).
- 12 Gavrilov, L. R. *Sov. Phys. Acoust.* 20, 16–18 (1974).
- 13 Hahn, G. M., Braun, J. & Har-Kedar, I. *Proc. natn. Acad. Sci. U.S.A.* 72, 937–940 (1975).
- 14 Hahn, G. M., Li, G. C. & Shiu, E. *Cancer Res.* 37, 761–764 (1977).
- 15 Lele, P. P. *Proc. Int. Symp. Cancer Therapy, Hyperthermia Radiation Am. Coll. Radiology*, 168–178 (1975).

Effect of human interferon preparations on lymphoblastogenesis in Down's syndrome

THE best characterised properties of human interferon, its antiviral (AV) and cell multiplication inhibitory (CMI) activities, are controlled, in an unexplained manner, by genes on chromosomes 21 (refs 1–4; 14–16). Human and animal interferons have various immunosuppressive effects^{5–10}, among them the inhibition *in vitro* of DNA synthesis in activated lymphocytes¹¹. Using mitogen- and antigen-stimulated lymphocytes from normal subjects (disomic 21) and others with Down's syndrome (trisomic 21), we have found that DNA synthesis is inhibited to a greater degree in the latter by both fibroblastoid and leukocyte interferons. We suggest that this property is also regulated by genes on chromosome 21.

Blood samples were obtained from ten subjects previously diagnosed with Down's syndrome, and from seven normal volunteers. The former were karyotyped, using quinacrine staining, to verify trisomy of chromosome 21. Lymphocytes were separated on a Ficoll-Hypaque gradient. Blast formation was stimulated either with phytohaemagglutinin (PHA) or in a one-way mixed lymphocyte reaction (MLR). Serial dilutions of two different interferon preparations, one from a poly(I):poly(C)-induced fibroblastoid cell line maintained in our laboratory³ and one from virus-induced leukocytes, were added at the beginning of culture. Serially diluted

Table 1 Sensitivity of trisomic 21 lymphocytes to interferon preparations

Trisomic 21 donor	Age	Fibroblast interferon		Leukocyte interferon	
		PHA	MLR	PHA	MLR
M.B.	16	4	1.5	4	1
N.C.	17	2	3	4	1.5
H.C.	28	25	32	1,000	—
J.S.	28	1	1	—	—
R.D.	32	2	4	850	6
D.B.	32	2	7	850	2
L.G.	9	11	—	30	4
S.N.	8	18	—	70	2
R.H.	14	5	5	85	7
D.L.	13	3	5	30	5

dilution of interferon preparation that produces 50% inhibition of DNA synthesis in trisomic 21 lymphocytes

Sensitivity =

dilution of interferon preparation that produces 50% inhibition of DNA synthesis in disomic 21 lymphocytes

Two trisomic 21 and one disomic 21 samples were tested on each occasion. Human fibroblast interferon was prepared by inducing human fibroblastoid cultures with poly (I)-poly (C). This interferon preparation was centrifuged at 100,000g for 1 h to remove any SV40 virus particles. Purified human leukocyte interferon¹³, prepared by Dr K. Cantell, was supplied by Dr M. Krim. Both interferon preparations were verified by Dr E. Havell as fibroblast and leukocyte interferon, respectively, by antibody neutralisation tests¹². Lymphocytes were cultured and stimulated with PHA or by allogeneic cells as described in Tables 2 and 3, and were treated with serially diluted interferon preparations for the duration of the experiments. Mock interferon preparations of the appropriate dilution were included with the experiments.

mock interferon, prepared in the same way as fibroblastoid interferon, but without inducer, was added to control cultures. DNA synthesis, an indicator of lymphoblastogenesis, was determined by measuring the incorporation of tritiated thymidine into the trichloroacetic acid (TCA)-precipitable cell fraction. To ensure that the interferon preparations did not alter the uptake of thymidine into the

Table 2 DNA synthesis in PHA-stimulated lymphocytes and unstimulated controls

Expt no.	Disomic 21		Trisomic 21	
	PHA	Control	PHA	Control
(1)	20,369	94	13,638 24,284	197 114
(2)	32,770	—	26,997 20,570	—
(3)	23,241	491	21,176 14,110	410 203
(4)	26,154	—	15,186 12,680	—
(5)	35,431	437	35,973 27,314	185 286

Figures are average c.p.m. from duplicate cultures.

Lymphocytes were separated from heparinised blood on a Ficoll-Hypaque gradient (Pharmacia) and washed three times in Ca^{2+} and Mg^{2+} free Hanks basic salt solution (Gibco). Duplicate sets of cultures were grown in round bottomed Microtest plates (Falcon) using RPMI 1640 medium (Gibco) containing 10% foetal calf serum (Flow), 1% chlorotetracycline and 0.1% PHA-P (Difco). Each well contained 5×10^4 cells in 0.2 ml. After 72 h, 0.01 ml of ^3H -thymidine (Schwartz-Mann, specific activity 1.9 Ci mmol⁻¹, 0.1 mCi ml⁻¹) was added. Four hours later, 20 μl of 20% sodium dodecyl sulphate was added to lyse the cells. The contents of each well were placed on glass fibre filter papers, washed three times in 5% TCA to precipitate the DNA and rinsed three times with 95% ethanol. After drying, the papers were placed in liquid scintillation vials with Spectrafluor (6 g PPO and 75 mg POPOP per l toluene, Amersham/Searle) and counted for 10 min, or to 1% error.

cell, the amount of radioactivity in the soluble fractions was also measured; no difference was found between the interferon-treated and untreated cultures.

In similar experiments, human interferon preparations were applied to PHA-stimulated murine spleen cells, while mouse interferon was added to human lymphocyte cultures. No species cross-reactivity was observed.

Trisomic 21 lymphocytes, whether stimulated by PHA or by allogeneic cells, were more sensitive than disomic 21 lymphocytes to the suppression of DNA synthesis by both interferon preparations (Table 1). The number of copies of chromosome 21 had no effect on the maximum levels of DNA synthesis obtained in the mock interferon-treated controls (Table 2). There was also no significant difference ($P > 0.5$) in maximum levels of DNA synthesis between the various donor combinations in the MLR, that is, it made no difference which genotype, trisomic 21 or disomic 21, was the stimulator or responder (Table 3). These results

Table 3 DNA synthesis in one-way mixed lymphocyte reactions

Expt no.	Responder $\times$ stimulator		T21 $\times$ T21
	T21 $\times$ D21	D21 $\times$ T21	
(1)	4,053 2,659	2,858 4,402	4,227 4,865
(2)	2,217 2,679	4,244 5,621	2,667 2,295
(3)	7,156 7,148	8,161 6,625	5,484 9,044
(4)	6,428 3,212	5,151 7,539	3,853 1,914
(5)	4,208 3,237	5,668 4,052	1,638 1,850

Figures are average c.p.m. from duplicate cultures. T21, trisomic 21; D21, disomic 21.

After separation, lymphocytes to be used as stimulator cells were treated for 30 min with mitomycin C ($25 \mu\text{g ml}^{-1}$), and washed twice in RPMI 1640 medium. Duplicate sets of cultures consisted of 5×10^4 stimulator and 5×10^4 responder cells in each well with 0.2 ml of complete medium. ^3H -thymidine (0.01 ml) was added after 96 h and cultures were terminated at 120 h.

indicate that the higher sensitivity of trisomic 21 lymphocytes to human interferon preparations is due to the presence of an extra copy of chromosome 21, and not to a changed sensitivity to antigenic or mitogenic stimuli.

The lymphocytes of one individual, J.S., who had been suspected to have Down's syndrome, showed a response to interferon similar to those of disomic 21 controls (Table 1); karyotyping revealed that J.S. had only two copies of chromosome 21. All trisomic 21 lymphocytes were more sensitive to the inhibitory effects of the human interferon preparations. This suggests that suppression of *in vitro* lymphoblastogenesis by human interferon, like its AV and CMI activities, acts through genes on chromosome 21. This implies that the various actions of human interferon applied to human cells are mediated through a common cellular pathway, the exact nature of which will probably not be known until the specific gene product(s) of chromosome 21 has been identified. Because interferons have not been purified and isolated conclusively, it remains to be seen whether the different effects of interferon preparations are due to one or more molecules. Until purification is achieved, and the respective antibodies prepared, it should be possible to assay human interferon preparations on the basis of their action through chromosome 21, in order to distinguish them from other macromolecules with similar actions, such as other lymphokines.

The developmental and metabolic abnormalities associated with Down's syndrome are probably due to the

presence of an extra chromosome and the resulting imbalance of gene products. Since trisomic 21 cells are hypersensitive to the CMI activities of human interferons, including the inhibition of lymphoblastogenesis, it is possible that some of the developmental and immunological defects associated with Down's syndrome are due to an abnormal interferon response.

We thank Dr C. Tan for karyotyping the Down's syndrome subjects, Dr E. A. Havell for characterising the interferon preparations immunologically and the MRC of Canada and IROI CA19721-01 for support. Y.H.T. is an NCI Scholar of Canada.

CLAIRE G. CUPPLES
Y. H. TAN

Department of Medical Biochemistry and Pediatrics,
University of Calgary,
Alberta, Canada T2N 1N4

Received 15 November 1976; accepted 8 March 1977.

- ¹ Tan, Y. H., Tischfield, J. & Ruddle, F. H. *J. exp. Med.* **137**, 317–330 (1973).
- ² Tan, Y. H. *Nature* **253**, 280–282 (1975).
- ³ Tan, Y. H. *Nature* **260**, 141–143 (1976).
- ⁴ Epstein, L. B. & Epstein, C. J. *Infect. Dis.* **133**, A56–A62 (1976).
- ⁵ Brodeur, B. R. & Merigan, T. C. *J. Immun.* **113**, 1319–1325 (1974).
- ⁶ Chester, T. J., Paucker, K. & Merigan, T. C. *Nature* **246**, 92–94 (1973).
- ⁷ DeMaeyer-Guignard, J., Cachard, A. & DeMaeyer, E. *Science* **190**, 574–576 (1975).
- ⁸ Donahoe, R. M. & Huang, K. Y. *Infect. Immun.* **13**, 1250–1257 (1976).
- ⁹ Lindahl, P., Leary, P. & Gresser, I. *Proc. natn. Acad. Sci. U.S.A.* **69**, 721–725 (1972).
- ¹⁰ Mobraaten, L. E., DeMaeyer, E. & DeMaeyer-Guignard, J. *Transplantation*, **16**, 415–420 (1973).
- ¹¹ Blomgren, H., Strander, H. & Cantell, K. *Scand. J. Immun.* **3**, 697–705 (1974).
- ¹² Havell, E. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 2185–2187 (1975).
- ¹³ Cantell, K., Hirvonen, S., Mogensen, K. E. & Pyhala, L. *Proc. Tissue Culture Ass. Workshop. The Production and Use of Interferon* (ed. Waymouth, C.) 35–38 (Tissue Culture Association, Rockville, Maryland, 1973).
- ¹⁴ Chany, C. *et al., Proc. natn. Acad. Sci. U.S.A.* **72**, 3129–3133 (1975).
- ¹⁵ Revel, M., Bash, D. & Ruddle, F. H. *Nature* **260**, 139–141 (1976).
- ¹⁶ De Clercq, E., Edy, V. G. & Cassiman, J.-J. *Nature* **264**, 249–251 (1976).

Dystrophic mice show age related muscle fibre and myelinated axon losses

ALTHOUGH many differences between age matched normal and dystrophic animals have been found, there have been few demonstrations of a time related quantitative change from normal to a characteristically dystrophic situation within the dystrophic strain itself. Here we report such a change—normal numbers of muscle fibres present in young dystrophic mice were rapidly lost, such that older animals showed the reduced number of muscle fibres characteristic of murine dystrophy. Although these losses began after a demonstrable loss of myelinated axons had occurred, it is not possible to say if the loss of muscle fibres was a result of the loss of nerve fibres.

The major practical problem in this study was that of obtaining animals of known dystrophic genotype of less than 3–4 weeks of age. The use of the C57BL/6J dy^{2J}/dy^{2J} strain of dystrophic¹ mice overcame this problem; affected animals occasionally breed and produce one or two litters. Thus, with animals of known dystrophic genotype, early developmental observations can be carried out.

A detailed investigation of the growth and development of normal 129ReJ $+/+$ and dystrophic 129ReJ dy/dy mouse muscles^{2,3} revealed a progressive loss of muscle fibres from the earliest age point at about 3 weeks. By about 6 weeks 40–50% of the muscle fibres, when compared with normal, had been lost from the respectively fast⁴ and slow-twitch⁴ extensor digitorum longus (EDL) and soleus muscles. Harris *et al.*⁵ and Bradley and Jenkinson⁶ have published conflicting reports on the number of myelinated axons in the nerves of 129ReJ dy/dy dystrophic mice. A 20% reduction of myelinated axons was reported by Law and Atwood⁷ in 129ReJ dy/dy dystrophic mouse soleus nerves. Bray and Aguayo⁸ observed a slight but probably not statistically significant reduction in the number of

myelinated axons in the nerve to the medial gastrocnemius muscle of 129B6F₁ dy/dy dystrophic mice.

Each of the studies mentioned above used a light microscope counting technique for axon numbers. We have found that these techniques often lead to an underestimation by 10–20% of the number of myelinated axons present when compared with counts obtained from low-power electron micrographs of the same nerve. We therefore used low-power electron micrographs of the nerves of male and female C57BL/6J $+/+$ normal and C57BL/6J dy^{2J}/dy^{2J}

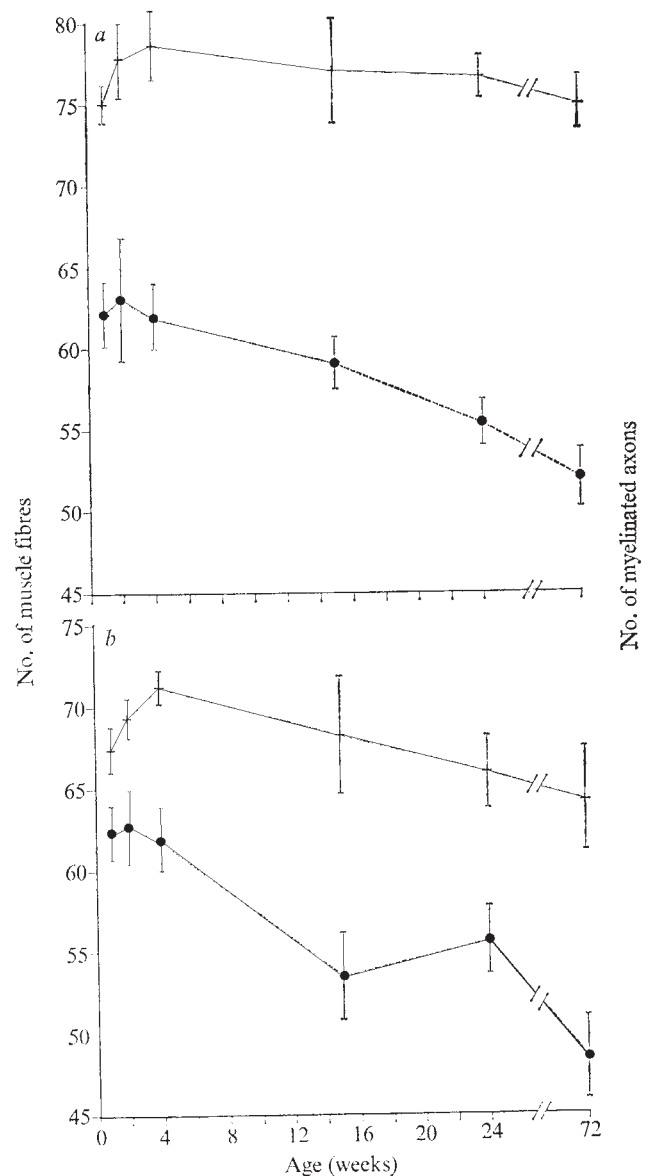


Fig. 1 *a* Muscle fibre counts from normal (+) and dystrophic (●) soleus muscles as a function of age. Each point represents the mean $\pm$ s.e. of at least 6 observations. The 1 week normal and dystrophic values were not statistically significantly different (Student's *t* test $P=0.05$ was used for all significance tests). At 15 and 72 weeks the percentage reductions in dystrophic fibre numbers were 37.4% and 33.8% respectively and the differences between the dystrophic and normal mean values at these age points were significantly different. *b*, Muscle fibre counts from normal (+) and dystrophic (●) extensor digitorum longus muscles. Each point represents the mean $\pm$ s.e. of at least 6 observations. The differences between the 1, 2 and 4 week normal and dystrophic values were not statistically significant. At 15 and 72 weeks the percentage reductions in dystrophic fibre numbers were 45.9% and 44.6% and the differences between the dystrophic and normal mean values at these age points were significantly different.

dystrophic animals aged 1–72 weeks. The nerves used in this study were those innervating the fast-twitch plantaris⁹ and slow-twitch⁴ soleus muscles. They both presented easily accessible, anatomically distinct, unifascicular portions,

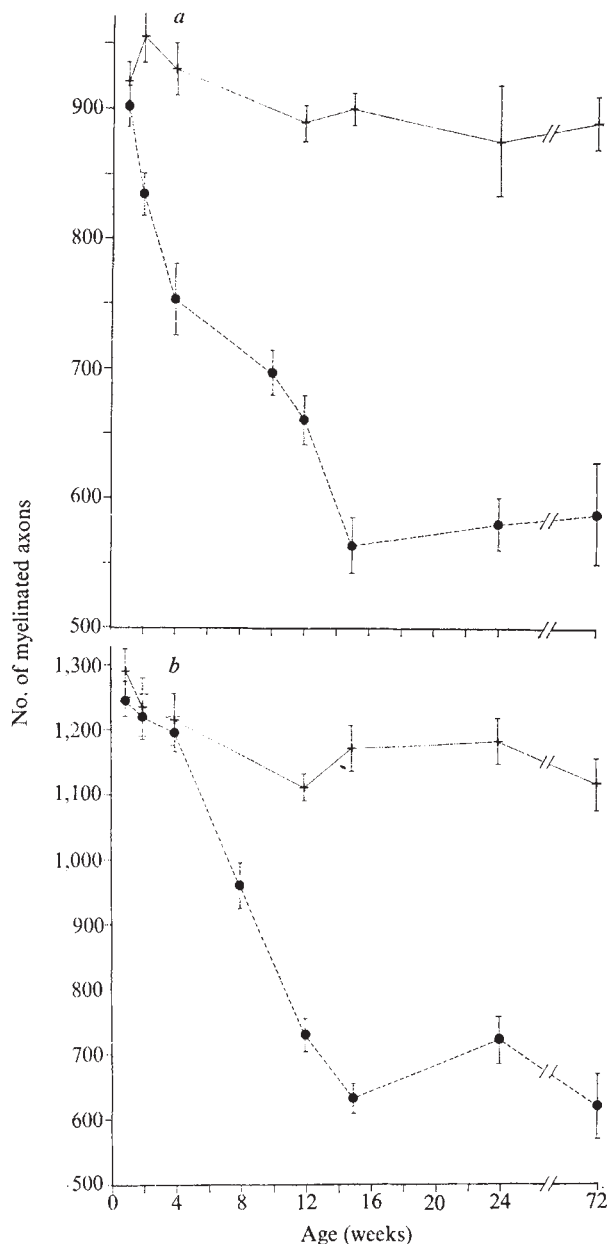


Fig. 2 *a* Myelinated axon counts from normal (+) and dystrophic (●) soleus nerves as a function of age. Each point represents the mean $\pm$ s.e. of at least 6 observations. The regression line for normal nerves was $y = -0.035x + 77.7$ and for dystrophic $y = -0.152x + 61.9$. The regression coefficient for normal nerves was not significantly different from zero whereas that for dystrophic nerves was. At 1 week the percentage reduction of dystrophic myelinated axons was 18.0% and at 72 weeks 30.8% and the difference between the normal and dystrophic mean values was statistically significant at every age point. The difference between the 1- and 72-week dystrophic values was significant and the percentage reduction in this case was 16.3%. *b*, Myelinated axon counts from normal (+) and dystrophic (●) plantaris nerves. Each point represents the mean $\pm$ s.e. of at least 6 observations. The regression line for normal values was $y = 0.074x + 69.2$ and for dystrophic values $y = 0.189x + 61.0$. The regression coefficient for normal nerves was not significantly different from zero whereas that for dystrophic nerves was. At 1 week the percentage reduction of dystrophic myelinated axons was 7.6% and at 72 weeks 24.3%. The differences between normal and dystrophic 1- and 2-week mean values were not statistically significant, at every other age point the differences were significant. The difference between the 1-week and 72-week dystrophic values was significant and the percentage reduction in this case was 21.9%.

from which transverse sections could be cut. After fixation for 2 h at 4 °C with a 3% solution of glutaraldehyde in 0.1 M sodium cacodylate buffer pH 7.4, the nerves were removed and placed in a 2% osmium tetroxide solution in 0.1 M sodium cacodylate buffer pH 7.4 at 4 °C for 1 h. After dehydration and embedding in Epon-Araldite, thin sections were cut for viewing with a Philips 200 electron microscope. The pennate structure of the plantaris muscle made it unsuitable for obtaining a count of the total number of muscle fibres from one transverse section. The EDL muscle, which histological examination had shown to be affected by the disease process, was therefore used as a fast-twitch muscle from which to obtain muscle fibre numbers. The slow-twitch soleus muscle was also studied. Muscles were fixed at their *in vivo* lengths with formol saline (NaCl 0.9%, formaldehyde 4.0%) then removed and treated for 1 h with 10% sulphuric acid to partly separate the fascicles¹⁰. Paraffin wax sections of 7 μ m thick were then cut from the belly of the whole muscle and stained with haematoxylin and eosin. Counting was carried out with the aid of a Leitz Prado microscope slide projector.

Figure 1 shows that rapid early reductions in the numbers of muscle fibres occurred in both soleus and EDL muscles. The percentage reductions in the EDL muscles were somewhat larger than in the soleus muscles and also began a little later. These changes were similar in magnitude to those observed by Rowe and Goldspink³ but, whereas the reduction in the 129ReJ *dy/dy* dystrophic mouse muscles were completed by about 6 weeks of age, those in the C57BL/6J *dy²³/dy²³* dystrophic mice were not completed until about 15 weeks of age. This again raises the question of two allelic genes having different phenotypic expression¹¹. The myelinated axon counts are shown in Fig. 2. In both dystrophic soleus and plantaris nerves the number of myelinated axons was always less than in their normal counterparts at each of the age points studied; these differences were significant. The percentage reduction in the myelinated axons of dystrophic soleus nerves was greater at every age point than the percentage reduction in the dystrophic plantaris nerves. At no age point was the percentage reduction in the nerve to the fast-twitch dystrophic plantaris muscles as great as that reported for the nerve³ to the fast-twitch^{12,13} tibialis anterior muscles. A progressive reduction in the number of myelinated axons also took place in both dystrophic plantaris and soleus nerves such that the 1- and 72-week values for each nerve were statistically significantly different. In addition, the regression coefficients were statistically significantly different from zero while those of normal values were not.

The areas of large unmyelinated axons, previously described in the C57BL/6J *dy²³/dy²³* dystrophic mice¹⁴, were not seen in the distal nerves we studied. We are also not yet able to say whether the myelinated axons that were lost were motor or sensory or both.

We thank Professor Harold Atwood for providing laboratory facilities, and Mrs I. Kwan and Mrs A. Welland for technical assistance. This work was supported by the Muscular Dystrophy Association of Canada.

A. MONTGOMERY
L. SWENARCHUK

Department of Zoology,
University of Toronto,
Toronto, Ontario M5S 1A1,
Canada

Received 30 December 1976; accepted 8 March 1977.

- Meier, M. & Southard, T. *Life Sci.* 9, 137–144 (1970).
- Rowe, R. W. D. & Goldspink, G. *J. Anat.* 104, 519–530 (1969).
- Rowe, R. W. D. & Goldspink, G. *J. Anat.* 104, 531–538 (1969).
- Close, R. *Nature* 206, 718–719 (1965).
- Harris, J. B., Wallace, C. & Wing, J. *J. Neurol. Sci.* 15, 245–249 (1972).
- Bradley, W. G. & Jenkinson, M. *J. Neurol. Sci.* 18, 227–247 (1973).
- Law, P. K. & Atwood, H. L. *Experientia* 30, 155–156 (1974).
- Bray, G. M. & Aguayo, A. J. *J. Neuropathol. exp. Neurol.* 34, 517–530 (1975).
- Binkhorst, R. A. & Voss, J. A. *Pflügers Archiv.* 296, 346–351 (1967).

- ¹⁰ Rayne, J. & Crawford, G. M. C. *J. Anat.* **119**, 34–357 (1975).
¹¹ Harris, J. B. & Montgomery, A. *Exptl Neurol.* **48**, 569–585 (1975).
¹² Harris, J. B. & Wilson, P. J. *Neurol. Neurosurg. Psychiat.* **34**, 512–520 (1971).
¹³ Montgomery, A. *J. neurol. Sci.* **26**, 401–423 (1975).
¹⁴ Bradley, W. G. & Jenkinson, M. *J. neurol. Sci.* **25**, 249–255 (1975).

Cs⁺ causes a voltage-dependent block of inward K currents in resting skeletal muscle fibres

WHEN frog skeletal muscle fibres are bathed in solutions containing Cs⁺ and K⁺ in the ratio 1:4,000, a reduction is observed in the size of inward K currents through the resting membrane. This effect is enhanced by an increase in either hyperpolarisation or external Cs⁺ concentration. It can be predicted from these findings that regenerative changes in membrane potential should be obtainable in fibres, in the presence of Cs⁺, that are hyperpolarised by means of a current electrode. Such responses are described in the last part of this report. In squid axon and frog node^{1,2}, internal Cs⁺ produces a voltage-dependent block of the delayed, outward K currents, though the ratio of Cs⁺ to K⁺ required for this effect is far greater than that used in the experiments reported here. A closer parallel can be drawn between our findings and those recently reported on the inward K currents in the starfish egg cell³.

Sartorius muscles of *Rana temporaria* were equilibrated in isotonic K₂SO₄ Ringer, containing 40 mM K₂SO₄, 8 mM CaSO₄, 1.08 mM Na₂HPO₄, 0.43 mM NaH₂PO₄ and 113 mM sucrose, pH 7.2. Additions of Cs₂SO₄ and K₂SO₄ were made by isotonic replacement of sucrose. Because K⁺ was the only permanent ion in these solutions, the resting potential was identical with the K equilibrium potential, V_K .

Membrane currents were measured using a three-electrode voltage-clamp technique⁴. Because with this method, membrane current is recorded as the potential difference between two intracellular microelectrodes, membrane current is expressed in mV. The membrane potential was held at V_K in all experiments.

Some properties of the resting K conductance are illustrated in Figs 1 and 2. The top of Fig. 1 shows records from two fibres, one (Fig. 1A) in 40 mM K₂SO₄, the other (Fig. 1B) in 40 mM K₂SO₄ plus 1.25 mM Cs₂SO₄. In the absence of Cs⁺ (Fig. 1A), the response to a step-wise hyperpolarisation to a membrane potential of -75 mV (upper record, a) was an inward current which turned on instantaneously and then declined slowly to a steady-state value. When the hyperpolarisation was increased (b), both instantaneous and steady-state currents were larger. The slow decline in K current has been attributed to a depletion of K⁺ adjacent to the membrane⁵. This depletion occurs because about 80% of the resting K conductance is located in the membrane of the transverse tubular system of muscle. During a maintained hyperpolarisation, K⁺ flowing into the sarcoplasm across the tubular wall is replaced partly by other ions from the bulk external medium, resulting in a fall in the tubular K⁺ concentration and consequent shift in V_K (ref. 6).

In the presence of 1.25 mM Cs₂SO₄ (Fig. 1B), both instantaneous and steady-state currents were smaller than for similar hyperpolarising voltages in the absence of Cs⁺. In addition, the current obtained for a large hyperpolarisation to -106 mV (lower record, b) was smaller than that recorded at only -81 mV (a). The ability of Cs⁺ to interfere with the inward movement of K⁺ must increase with hyperpolarisation. It is not clear from these records in Cs⁺ whether the time-dependent decline in current arises from a K depletion, a time-dependent block by Cs⁺, or both.

Mean instantaneous current-voltage relations for a number of fibres in solutions containing 0.01–1.25 mM

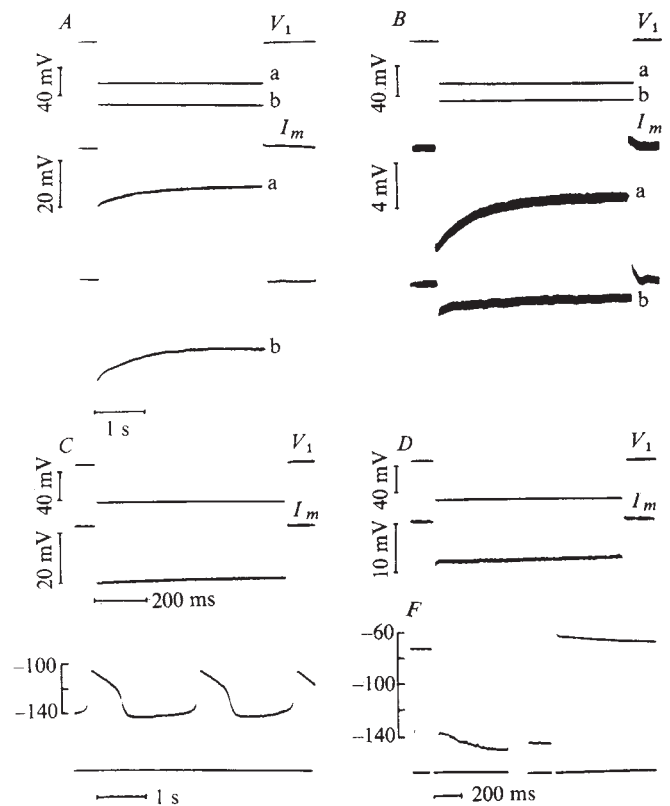


Fig. 1 A, records of membrane potential (above) and membrane current (below) from a fibre in 40 mM K₂SO₄ Ringer. Fibre resting potential (RP), -14 mV; holding potential (HP), -14 mV; T, 17.2 °C. B, as (A) except that the solution also contained 1.25 mM Cs₂SO₄. Fibre RP, -15 mV; HP, -15 mV; T, 17.2 °C. Time scale as in (A). C, record of membrane potential (above) and membrane current (below) from a fibre in 95 mM K₂SO₄ Ringer. RP, +3 mV; HP, +3 mV; T, 4 °C. D, as (C) except that the solution also contained 1.25 mM Cs₂SO₄. Fibre RP, +1 mV; HP, +1 mV; T, 3.8 °C. Time scale as in (C). E–F, records were made using the three-electrode voltage-clamp technique, where membrane current is recorded as the potential difference between two impaling microelectrodes, recording membrane potentials V_1 and V_2 , at distances l and $2l$ from the pelvic end of the fibre. The membrane current is then given by the approximation⁴

$$I_m = \frac{a(V_2 - V_1)}{3l^2 R_i} \text{ A cm}^{-2} \quad (1)$$

where a is the fibre radius and R_i is the resistivity of the sarcoplasm. The value of l was 330 μ m. The calibration on the current records is in millivolts since the records are of $V_2 - V_1$. If a is assumed to be 40 μ m, and R_i to be 180 Ω cm at 17–18 °C and 290 Ω cm at 3–4 °C, 1 mV equals 6.8 μ A cm⁻² at 17–18 °C and 4.2 μ A cm⁻² at 3–4 °C. E, record of membrane potential (above) and current flowing through an impaling microelectrode (below) made in constant current conditions with two impaling electrodes. The fibre is in a solution containing 40 mM K₂SO₄ and 0.25 mM Cs₂SO₄. Vertical scale: membrane potential (mV). F, records made as in (E), except that the fibre is in a solution containing 40 mM K₂SO₄ and 1.25 mM Cs₂SO₄.

Cs₂SO₄ are given in Fig. 2. In the absence of Cs⁺ (solid curve), the size of the K current depends on the direction in which K⁺ moves across the membrane. When the K-driving force (membrane potential, V_i , minus V_K) is negative, relatively large inward currents can be recorded. In contrast, when V_i minus V_K is positive and the K current is outward, currents are always small. The current-voltage relation is therefore that of an incoming rectifier⁷.

Addition of Cs⁺ (Fig. 2) did not affect outward K currents. Similarly, when the inward driving force was small, inward currents also resembled those in the control Ringer. As V_i minus V_K was made more negative, however, the inward currents reached a maximum, then declined with further

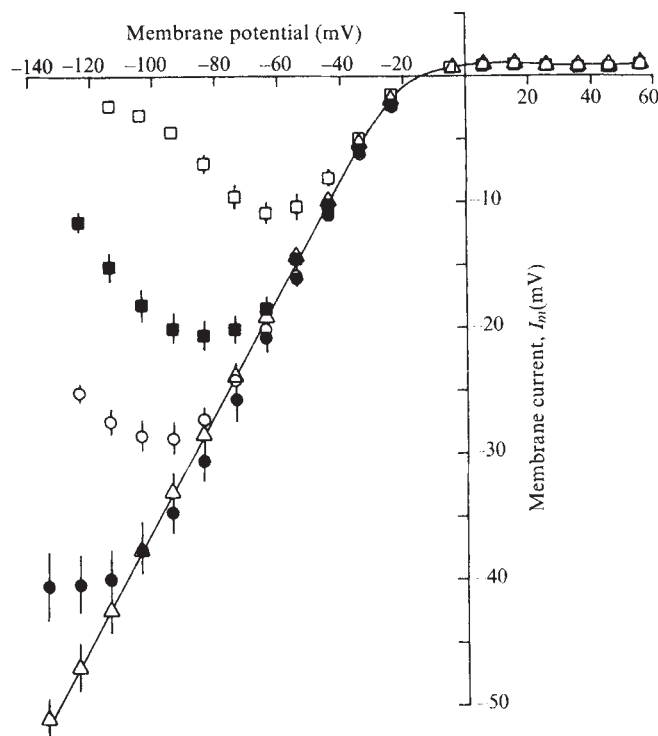


Fig. 2 Instantaneous current-voltage relations recorded with a three-electrode voltage clamp in muscle fibres. Abscissa, membrane potential, V_i (mV). Ordinate, membrane current, I_m (mV). In the solutions used the membrane current was carried by K^+ . The instantaneous current was obtained by exponential extrapolation of the time-dependent current (Fig. 1A and B) to the beginning of the voltage step. Each experimental point is the mean of results from six to eight fibres. Vertical lines, s.e.m. T , 17–18°C. Fibres were held at the resting potential (which was equal to V_K). RP, -14.13 ± 0.24 mV ($n = 48$) in 40 mM K_2SO_4 , and -13.92 ± 0.37 mV ($n = 24$) in 40 mM K_2SO_4 plus 1.25 mM Cs_2SO_4 (mean $\pm$ s.e.m. of all penetrations, $n =$ no. of observations). Medium was 40 mM K_2SO_4 Ringer plus: Δ , no Cs_2SO_4 ; $\bullet$, 0.01 mM Cs_2SO_4 ; $\circ$, 0.05 mM Cs_2SO_4 ; $\blacksquare$, 0.25 mM Cs_2SO_4 ; $\square$, 1.25 mM Cs_2SO_4 .

hyperpolarisation. It can be concluded from these results that the block of instantaneous K currents by Cs^+ is a function of both the K -driving force and the external Cs^+ concentration.

To test whether the blockage by Cs^+ was purely instantaneous, currents were recorded in fibres in a 95 mM K_2SO_4 Ringer, in the cold. In these conditions the complications introduced by K depletion are much reduced, and the currents show very little time dependence in the control Ringer (Fig. 1C). Nor is there any noticeable time dependence (apart from the initial capacity transient) in the presence of 1.25 mM Cs_2SO_4 (Fig. 1D), which is consistent with the notion that the blockage is instantaneous, and has no time-dependent component.

The measurement of currents at extreme hyperpolarisation is complicated by contractile repriming and membrane breakdown. In solutions containing 40 mM K_2SO_4 plus 0.25 mM Cs_2SO_4 , however, instantaneous currents reached a minimum at -155 mV, then increased gradually with further hyperpolarisation. The current-voltage relation determined in the presence of Cs^+ for membrane potentials between -14 mV and -200 mV can therefore be intercepted at three membrane potentials by a line (drawn parallel to the voltage axis) corresponding to a small inward current. Because two of these potentials will be stable and one unstable, theoretically⁸ it should be possible to record regenerative responses from the unclamped, hyperpolarised membrane.

Muscle fibres were impaled near their mid-points by two electrodes, one to record the membrane potential, and the

other to pass current. When the hyperpolarisation produced by a steady inward current exceeded -115 mV in the presence of 0.25 mM Cs_2SO_4 , the membrane potential developed spontaneous, stable oscillations (Fig. 1E), with distinct thresholds for both a depolarisation and a repolarisation phase, reminiscent of the cardiac action potential. The depolarisation phase of the cycle is probably due to a regenerative increase in K permeability, produced by the potential-dependent dissociation of Cs^+ from the membrane. The potential slowly becomes more negative again as K depletion occurs, until a threshold is reached at which Cs^+ can again bind to the membrane, resulting in a regenerative decrease in K permeability and consequent rehyperpolarisation. That the thresholds and the peak depolarisation do not correspond correctly to predicted points on the current-voltage curve (Fig. 2) is in part due to the difference between the uniform and point methods of membrane polarisation⁸. In addition, the current-voltage relationship varies with time; this is also the reason why the responses are cyclical and not maintained.

Maintained, reversible transitions between two stable potentials (about -75 mV and -140 mV) were produced by brief current pulses superimposed on a steady inward current in 40 mM K_2SO_4 plus 1.25 mM Cs_2SO_4 (Fig. 1F). In these conditions, however, cyclical responses were never seen.

The ability to record two types of regenerative responses in hyperpolarised fibres also supports the conclusion that the blockage of inward currents by Cs^+ depends on both the membrane potential and the Cs^+ concentration. The stability of the responses indicates that the potential-dependent increase in instantaneous current at extreme hyperpolarisations is due to a recovery from Cs^+ blockage, and is not an artefact of membrane breakdown.

We thank Dr N. B. Standen for reading the manuscript. L.A.G. is supported by an MRC Studentship.

LOUISE A. GAY
P. R. STANFIELD

Department of Physiology,
University of Leicester,
Leicester, UK

Received 28 February; accepted 11 March 1977.

- 1 Armstrong, C. M. in *Membranes, A Series of Advances* 3 (ed. Eisenman, G.) 325–358 (Dekker, New York, 1975).
- 2 Hille, B. in *Membranes, A Series of Advances* 3 (ed. Eisenman, G.) 255–323 (Dekker, New York, 1975).
- 3 Hagiwara, S., Miyazaki, S. & Rosenthal, N. P. *J. gen. Physiol.* 67, 621–638 (1976).
- 4 Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol.* 208, 607–644 (1970).
- 5 Adrian, R. H. & Freygang, W. H. *J. Physiol.* 163, 61–103 (1962).
- 6 Almers, W. *J. Physiol.* 225, 33–56 (1972).
- 7 Adrian, R. H. *Prog. Biophys. molec. Biol.* 19, 339–369 (1969).
- 8 Noble, D. & Hall, A. E. *Biophys. J.* 3, 261–274 (1963).

Is hyperosmotic neurosecretion from motor nerve endings a calcium-dependent process?

SPONTANEOUS liberation of neurotransmitter quanta is strongly affected by the osmotic pressure of the extracellular fluid. Elevation of the osmolarity by 20–30% increases the rate of release from motor nerve endings by more than one order of magnitude^{1,2}. In this respect the neuromuscular junction resembles some other secretory systems^{3–5}. The mechanism of this hyperosmotic neurosecretion is not yet understood; extracellular calcium ions are not directly responsible, since this effect can be produced in their absence⁶. Recently, it has been suggested that the liberation of neurotransmitter is regulated by the intracellular concentration of free calcium ions^{6–8}. We have therefore examined the hypothesis that hyperosmotic neurosecretion originates from an increase in internal calcium concentration ($[Ca]_{in}$). At the frog neuromuscular synapse however, it is impossible at present to estimate directly free $[Ca]_{in}$; hence we used an indirect technique,

which is based on two assumptions: first, the frequency of the miniature endplate potentials (m.e.p.p.s.) reflects free $[Ca]_{in}$. Second, the movement of calcium ions across the presynaptic membrane is governed by the electrochemical gradient, and by the calcium conductance (g_{Ca}). If hyperosmotic neurosecretion is caused by an increase in $[Ca]_{in}$, then increasing g_{Ca} under reversed electrochemical gradient for the calcium should cause a reduction in the effect of hyperosmotic stress on transmitter release. We report that hyperosmotic neurosecretion is dependent on $[Ca]_{in}$.

The experiments were performed on the sartorius nerve muscle preparation of the frog *Rana pipiens*, using conventional methods for intracellular recording. The standard bathing Ringer solution contained 116 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$ (pH adjusted to 7.0 by titration). The extracellular calcium concentration was either 0.4 mM or less than 10^{-8} M when 1 mM EGTA was added⁹.

Figure 1 illustrates the main experimental observation. A preparation was bathed in a standard Ringer solution with 1 mM EGTA. At A, 50 mM sucrose were added to the medium. This hyperosmotic solution produced a marked increase in the rate of spontaneous transmitter release (miniature end plate potentials frequency). At B, a second perturbation was introduced. The concentration of $[K]_0$ was increased from 2 to 20 mM by iso-osmotic substitution for sodium. This increase in $[K]_0$ is known to produce a depolarisation and an increase in the calcium conductance of nerve membrane¹⁰⁻¹². In normal circumstances, this increase in calcium permeability is expected to cause a calcium influx along the electrochemical gradient for calcium and an increase in transmitter release. At B however, due to the presence of 1 mM EGTA, there is a reversed electrochemical gradient and the increase in calcium permeability presumably causes an efflux of calcium, leading to a decrease in transmitter release. Restoring the inward electrochemical gradient for calcium in C produces a marked increase in m.e.p.p. frequency.

Although extracellular calcium is not necessary to produce hyperosmotic neurosecretion, it has nevertheless an effect on the magnitude of the response. This is illustrated in Fig. 2, where hyperosmotic neurosecretion was induced at A, by adding 50 mM sucrose to standard Ringer solution with 0.4 mM $CaCl_2$. At B the $[Ca]_{out}$

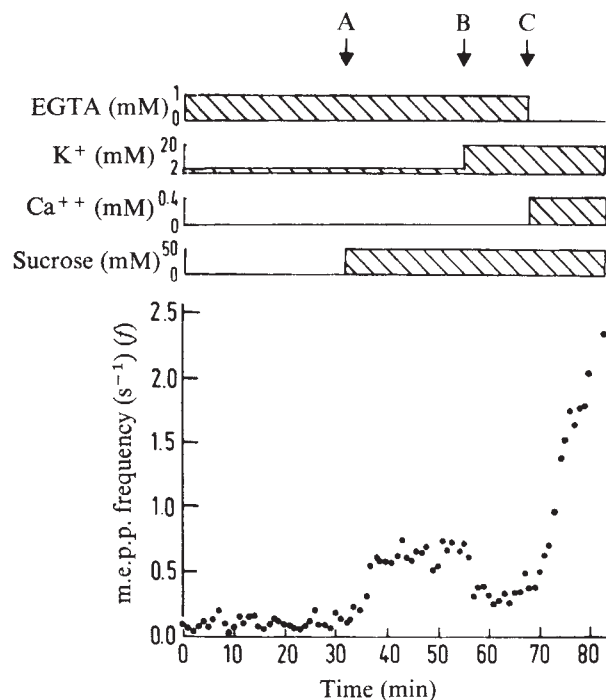


Fig. 1 Hyperosmotic neurosecretion at the frog neuromuscular synapse. A micropipette was placed near the endplate region and the frequency (f) of the spontaneous miniature endplate potentials was monitored at various compositions of the extracellular medium. The control solution contained 116 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, 1 mM EGTA, with free $[Ca]_0$ of less than 10^{-8} M (with presumably a reversed gradient for calcium ions across the presynaptic membrane). At A the osmolarity of the medium was increased from 225 to 275 mosM, by adding 50 mM sucrose. It is clear that hyperosmotic neurosecretion can be obtained in the virtual absence of calcium ions. At B $[K]_0$ was increased to 20 mM. In contrast to the behaviour in a normal calcium gradient, increase in $[K]_0$ in reversed gradient caused a decrease in f . Restoring an inward calcium gradient in C caused a marked increase in f . Room temperature was approximately 20 °C. The entire experiment was performed at the same fibre, with continuous perfusion of the bathing fluid at an average rate of approximately 2 ml min^{-1} .

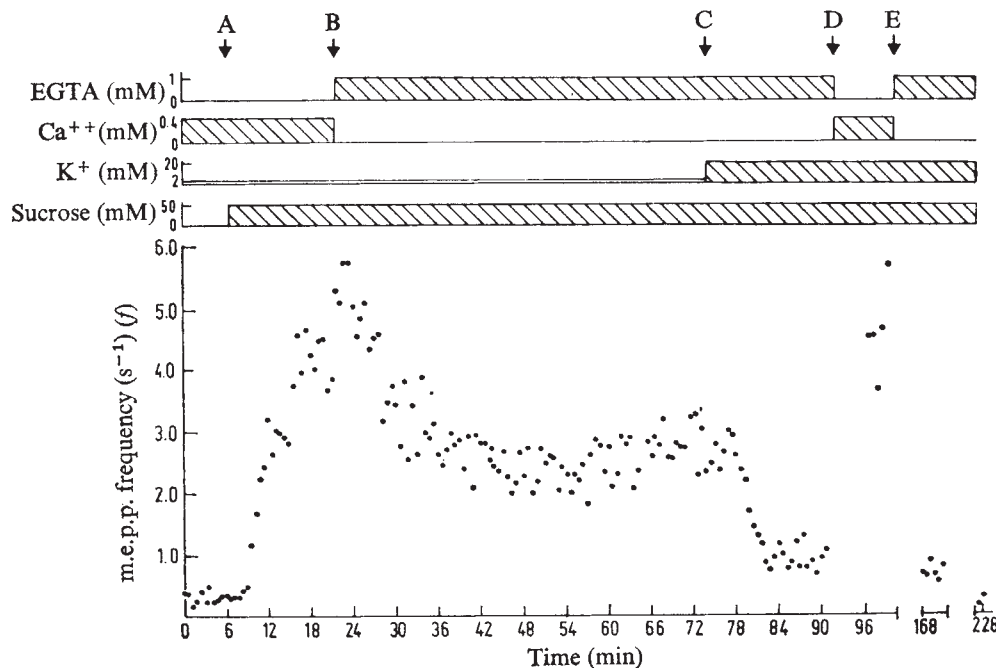


Fig. 2 Factors affecting hyperosmotic neurosecretion at the frog neuromuscular junction. The experiment was started in conditions similar to those in Fig. 1, except that the initial $[Ca]_0$ was 0.4 mM (and no EGTA). Hyperosmotic neurosecretion was produced (at A) by 50 mM sucrose. At B $[Ca]_0$ was reduced to less than 10^{-8} M by adding 1 mM EGTA and omitting $CaCl_2$. Comparison of Fig. 1 and Fig. 2 shows that although $[Ca]_0$ is not necessary for hyperosmotic neurosecretion, it nevertheless influences the magnitude of the response; in the absence of $[Ca]_0$ hyperosmotic neurosecretion is reduced to 40–80%. Addition of $[K]_0$ at C caused again a very substantial decrease in m.e.p.p. frequency. In D and E $CaCl_2$ (0.4 mM) was reintroduced and removed respectively, producing the expected changes in f (see text). Room temperature at 20 °C and continuous perfusion with an average rate of approximately 2 ml min^{-1} .

was reduced to less than 10^{-8} M by omitting calcium and adding 1 mM EGTA; this reduced the m.e.p.p. frequency to approximately 50%. Here again a further decrease was obtained by depolarising the membrane by 20 mM KCl (C).

These experiments suggest that hyperosmotic neurosecretion, similar to other processes which modulate transmitter release^{13,14}, depends on changes in $[Ca]_{in}$. The source for the increase in $[Ca]_{in}$ is not yet clear. One possible origin is the mitochondrion¹⁵, since it has been shown that increase in the osmolarity around liver mitochondria inhibits calcium uptake¹⁶.

We thank Sir Bernard Katz for helpful comments, the Director and staff of Cold Spring Harbor Laboratory for generous hospitality, and Mr Bruce DeTroy, Mr Robert Yaffe, Miss Ester Eilander and Mr Shlomo Ben Yona for technical assistance. This work was supported by grants from the Sloan Foundation and the US-Israel Binational Science Foundation.

Y. SHIMONI
E. ALNAES*
R. RAHAMIMOFF

Department of Physiology,
Hebrew University-Hadassah Medical School,
P.O. Box 1172, Jerusalem, Israel, and
Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York 11724

Received 18 January; accepted 11 March 1977.

*Present address: Institute of Neurophysiology, University of Oslo, Oslo 1 Norway.

- 1 Fatt, P. & Katz, B. *J. Physiol., Lond.* **117**, 109–128 (1952)
- 2 Furshpan, E. J. *J. Physiol., Lond.* **134**, 689–699 (1956).
- 3 Hayward, J. N. & Vincent, J. D. *J. Physiol., Lond.* **210**, 947–972 (1970).
- 4 Moses, A. M., Miller, M. & Streeter, D. M. *P. Metabolism* **25**, 697–721 (1976).
- 5 Hays, R. M. *New Engl. J. Med.* **295**, 657–665 (1976).
- 6 Blich, Z., Glagoleva, I. M., Lieberman, E. A. & Nenashv, V. A. *J. Physiol., Lond.* **199**, 11–35 (1968).
- 7 Rahamimoff, R. & Alnaes, E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3613–3616 (1973).
- 8 Miledi, R. *Proc. R. Soc. B.* **183**, 421–425 (1973).
- 9 Baker, P. F., Hodgkin, A. L. & Ridgway, E. B. *J. Physiol., Lond.* **218**, 709–755 (1971).
- 10 Liley, A. W. *J. Physiol., Lond.* **134**, 427–433 (1956).
- 11 Baker, P. F., Meves, M. & Ridgway, E. B. *J. Physiol., Lond.* **231**, 527–548 (1973).
- 12 Cooke, J. D. & Quastel, D. M. *J. Physiol., Lond.* **228**, 435–458 (1973).
- 13 Katz, B. & Miledi, R. *J. Physiol., Lond.* **203**, 459–487 (1969).
- 14 Alnaes, E. & Rahamimoff, R. *J. Physiol., Lond.* **238**, 285–306 (1975).
- 15 Rahamimoff, R. *et al. Cold Spring Harb. Symp. quant. Biol.* **40**, 107–116 (1976).
- 16 Scarpa, A. & Azzone, G. F. *J. Biol. Chem.* **243**, 5132–5138 (1968).

Sodium and calcium action potential in pituitary cells

SEVERAL endocrine cells or their neoplastic derivatives generate action potentials similar to those seen in neurones¹, and in the adrenal chromaffin cell such regenerative potentials depend primarily on a sodium mechanism^{2–4}. Kidokoro has described action potentials in the GH3 rat pituitary cell line which seem to depend on a calcium mechanism⁵. We have re-investigated the action potential in GH3 cells and found that it results from combined Na and Ca mechanisms in physiological conditions. In addition, we have recorded similar electrical excitability from two human pituitary tumours grown *in vitro*.

GH3 cells, provided through the courtesy of Dr Gordon Sato, were dissociated mechanically and grown in 60-mm plastic tissue culture dishes (Falcon) in Dulbecco's MEM with heat-activated serum (5% foetal calf serum (FCS), 10% horse), penicillin (100 U ml⁻¹) and streptomycin (100 µg ml⁻¹). Human pituitary cells were obtained from surgical specimens of an acidophilic adenoma in an acromegalic 51-year-old woman and of a chromophobe adenoma which recurred after surgery and irradiation in a 53-year-old man without endocrine symptomatology. Pieces of tissue were dissociated mechanically or by treatment with 0.027% trypsin for 2.5 h followed by trituration. Dissociated cells were plated in 35-mm Falcon tissue culture dishes (1–6 × 10⁵ cells per dish) which had been coated with gelatine or a

primary culture of cytosine arabinoside-inhibited non-neuronal rat embryo cells. Human pituitary cells were maintained in McCoy's 5A with 20% FCS or MEM with 10% FCS, 200 mg% glucose and antibiotics.

Cells were penetrated with 3.5 M potassium acetate-filled microelectrodes (40–100 MΩ) under guidance of direct vision by phase contrast microscopic control as described previously⁴, except that self-filling microelectrodes were used sometimes. For recording, the cells were bathed in physiological saline consisting of 150 mM Na, 154 mM Cl, 3.5 mM K, 1.8 mM Ca, 1.0 mM Mg and 2.6 mM phosphate (pH 7.3). Occasionally, 2.5 mM Tris buffer was substituted for phosphate buffer.

In most stable and active penetrated GH3 cells, the resting potential reached between –40 and –67 mV, usually after improving from an initially less negative value. The input resistance also increased substantially to values between 40 and 525 MΩ (mean ± s.e. for 13 active cells of 15 with V_m equal to or more negative than –40 mV, 227 ± 38 MΩ). Stimulation of the cells by an intracellular depolarising pulse produced short duration (duration at half height, 2 ms) action potentials (Fig. 1a–c). These occasionally reached the null potential or showed a small overshoot (Fig. 1a). Multiple action potentials could occur in response to a single prolonged stimulus. In cells of 15 with V_m equal to or more negative than injured by impalement, background hyperpolarisation was usually necessary before stimulation could produce an action potential.

Human pituitary cells maintained in culture appeared as small round or spindle shaped cells (Fig. 2a). Cells from both tumours were able to generate action potentials similar to those seen in GH3 (Fig. 2b).

To determine the ionic basis of the electrical activity, cells were first penetrated in control saline, then in

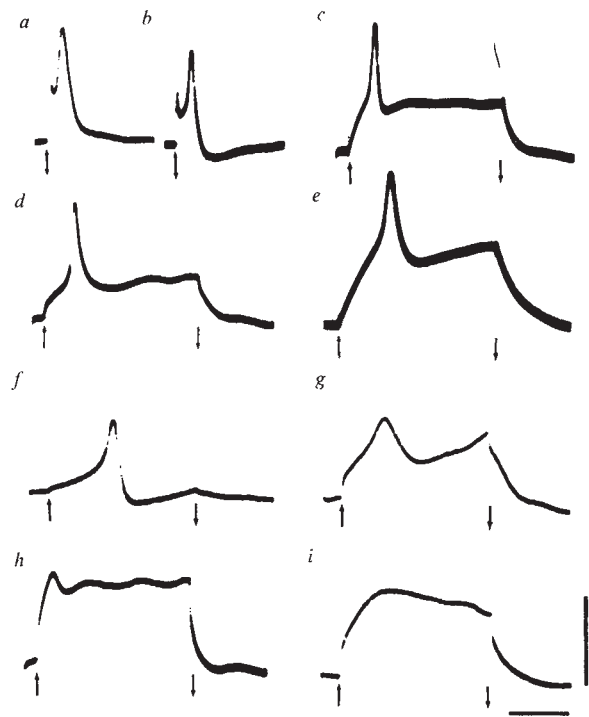


Fig. 1 Action potentials in GH3 cells. *a* and *b*, Control medium, 0.5-ms stimulating pulse at double-headed arrows; *c*, control medium, 50-ms stimulating pulses; *d*, Ca-free; *e*, Mn²⁺; *f*, TTX; *g*, Na-free; *h*, Ca-free and TTX; *i*, Mn²⁺ and TTX. Membrane potentials *a*–*i*: 52, 56*, 59, 74*, 80*, 46, 54, 80*, 55 mV. (* is V_m with background hyperpolarising current and is only approximate for (*d*) and (*h*) due to bridge imbalance.) Stimulus currents *c*–*i*: 0.5, 0.4, 1.3, 0.2, 0.6, 0.6 nA. Calibration bars in (*i*) are 50 mV and 20 ms in all traces.

Table 1 Results of three sets of experiments

a	Control	Ca-free	Ca-free + TTX	Ca-free
No. of active cells/Total no. of cells	14/15	9/9	0/22	10/12
Resting potential (mean $\pm$ s.e.)	22 $\pm$ 3 mV	12 $\pm$ 1	17 $\pm$ 1	16 $\pm$ 2
(range)	(11–50)	(10–15)	(11–37)	(10–35)
b	Control	Mn	Mn + TTX	Mn
No. of active cells/Total no. of cells	6/6	7/10	0/16	13/16
Resting potential	32 $\pm$ 7	23 $\pm$ 3	24 $\pm$ 2	25 $\pm$ 2
(range)	(16–55)	(12–41)	(12–42)	(16–38)
c	Control	TTX	TTX + Mn	TTX
No. of active cells/Total no. of cells	10/12	8/14	0/12	7/16
Resting potential	22 $\pm$ 3	29 $\pm$ 4	26 $\pm$ 6	25 $\pm$ 4
(range)	(10–40)	(10–70)	(10–66)	(12–56)

solutions which either omitted Ca or Na (after four washes), or contained inhibitors of Na or Ca spike mechanisms (tetrodotoxin (TTX) at 10^{-6} g ml $^{-1}$ to inhibit Na spikes, and Mn $^{2+}$ at 5 mM or Co $^{2+}$ at 7.5 mM to inhibit Ca spikes). Chelating agents were not used. Recordings were then made again in control solution. The proportion of cells with maintained resting potentials equal to or more negative than -10 mV which could generate action potentials was determined. The results of three sets of experiments are shown in Table 1, which includes the mean resting potentials ($\pm$ s.e.) and the range of the resting potentials recorded in each condition.

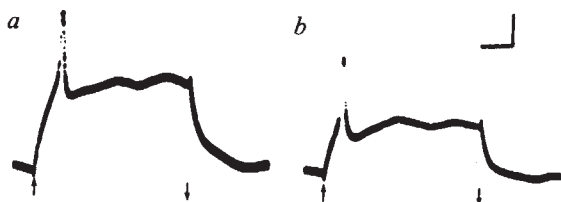


Fig. 2 Human pituitary cells. *a*, Cultured chromophobe adenoma cells seen with phase contrast microscopy after glutaraldehyde fixation. Calibration bar, 50 μ m. *b*, Action potential in control medium from cell from acidophil adenoma. $V_m = 68$ mV, stimulus current = 0.5 nA. *c*, Action potential in Ca-free medium from cell from chromophobe adenoma. $V_m = 80$ mV, stimulus current = 0.7 nA. Calibration = 20 mV and 10 ms.

When Ca was omitted from the medium, action potentials remained frequent and had rates of rise comparable with those in the control (Fig. 1*d*), in spite of lower resting potentials and input resistances. Action potentials were also demonstrated in the presence of 55 mM Mn $^{2+}$ (Fig. 1*e*). To confirm that the spikes recorded in the absence of Ca or in the presence of Mn $^{2+}$ were Na spikes, TTX was added to the Ca-free and Mn $^{2+}$ media. Activity disappeared within minutes (Fig. 1*h*), but could easily be recovered in both cases once the TTX was washed out (Table 1*a* and *b*).

When TTX was added to control saline (1.8 mM Ca), substantial action potentials were still present (Fig. 1*f*). These had lower maximum rates of rise (15 ± 2 V s $^{-1}$ ($n=6$) compared with 58 ± 7 V s $^{-1}$ ($n=13$) in control (active cells with V_m equal to or more negative than -40 mV in both cases)). Rates of rise were measured from photographic records of swept-out action potentials in cells which had been activated maximally by background hyperpolarisation. When Na was replaced in the medium by equiosmolar Tris, action potentials were still found, but with more difficulty (Fig. 1*g*). Spikes were more readily found when Ca was increased to 10 mM. We believe that the difficulty in Na-free conditions was probably due to nonspecific deterioration of the cells. To confirm that the spikes recorded in TTX were Ca spikes, Mn $^{2+}$ was added to the TTX medium. Action potentials promptly disappeared (Table 1*c*), although slight irregularity occasionally remained in the rising phase of the charging curve (Fig. 1*h*).

These experiments demonstrated that the rising phase of the action potential was generated by both a Na and a Ca

mechanism in physiological ionic conditions. The repolarising phase of the action potential could often be seen to undershoot the resting potential (Fig. 1*b*), and reversed at between -50 and -70 mV. If tetraethylammonium chloride (10 mM), an agent known to inhibit the active post-spike K conductance, was added to the medium, the action potential developed a plateau which could last up to 300 ms before repolarisation. It is likely, as in other excitable cells, that the rapid repolarisation of the membrane which terminates the spike is due to an increased K conductance.

The human pituitary tumour studies were limited by the small amounts of tissue in surgical specimens. The cells remained electrically excitable in Ca-free medium (Fig. 2*c*), and in the presence of Co $^{2+}$ (in the one tumour so studied), indicating that a large part of the spike was probably due to Na. The recordings in TTX, however, did not enable us to make a definite statement about the presence or absence of a major Ca component of the spike.

It has been established in various neuronal and endocrine tissues that output of transmitter or hormone is coupled with influx of Ca (ref. 6). In this regard it would seem useful for endocrine cells to utilise Ca-spike mechanisms to enhance excitation-secretion coupling. There is evidence for large Ca-spikes in GH3 cells and pancreatic B cells 7,8 but not in adrenal medulla 2,3 , pheochromocytoma or bronchial carcinoids 1 . In those endocrine cells without large Ca spikes it is likely that Ca entry could occur during depolarisation due to ACh (adrenal medulla) or other stimulatory agents, or during the Na action potential 3,9 . The function of Na action potentials in cells with large Ca spikes is less clear. They could help to spread excitation among electrically coupled cells, or they could act in some way as modulators of the Ca mechanism.

The data presented here apply to neoplastic derivatives of anterior pituitary cells. Action potentials have not been reported in normal anterior pituitary cells, but Davis and Hadley 10 have recorded extracellular spikes from cells in the pars intermedia of the frog. Preliminary recordings from unidentified rat and gerbil anterior pituitary cells *in vitro* in our laboratory have demonstrated action potentials. The detailed physiology of these responses is yet to be determined.

M.A.D. was supported by a Research Career Development Award from the NINCDS and a grant from the Esther A. and Joseph Klingenstein Foundation. A.T. was supported in part by grants from the NIH. We thank Dr Martin Posner, Ms Sara Vasquez and Ms Isabel Simmons for help in obtaining and maintaining cells in culture.

BERNARD BIALES
MARC A. DICHTER
ARTHUR TISCHLER

Departments of Neurology and Pathology,
Beth Israel Hospital/Harvard Medical School,
Boston, Massachusetts 02215

Received 30 December 1976; accepted 15 March 1977.

1 Tischler, A., Dichter, M., Biales, B., DeLellis, R. & Wolfe, H. *Science* 192 902–904 (1976).

- ² Biales, B., Dichter, M. & Tischler, A. *Neurosci. Abstr.* 1, 460 (1975).
³ Brandt, B., Hagiwara, S., Kidokoro, Y. & Miyazaki, S. *J. Physiol., Lond.* 263, 417-439 (1976).
⁴ Biales, B., Dichter, M. & Tischler, A. *J. Physiol., Lond.* 252, 743-753 (1976).
⁵ Kidokoro, Y. *Nature* 258, 741-742 (1975).
⁶ Rubin, R. *Calcium and the Secretory Process* (Plenum, New York, 1974).
⁷ Dean, P. & Mathews, E. *J. Physiol., Lond.* 210, 265-275 (1970).
⁸ Meissner, H. & Schmelz, H. *Pflug. Arch.* 351, 195-206 (1974).
⁹ Baker, P., Hodgkin, A. & Ridgway, E. *J. Physiol., Lond.* 218, 709-755 (1971).
¹⁰ Davis, M. & Hadley, M. *Nature* 261, 422-423 (1976).

Caffeine restores feeding response to 2-deoxy-D-glucose in 6-hydroxydopamine-treated rats

A LARGE portion of the central catecholaminergic nerve terminals of the rat are destroyed by administering 6-hydroxydopamine (6-HDA) via the cerebrospinal fluid¹⁻³. Animals lesioned in this way often appear normal, yet show many subtle behavioural abnormalities⁴. We have been examining one example of this phenomenon, the failure of 6-HDA-lesioned rats to increase food intake when given a systemic injection of 2-deoxy-D-glucose (2-DG) (refs 5, 6). This glucose analogue seems to elicit feeding in intact rats due to its inhibition of glycolysis in cerebral chemoreceptor cells^{7,8}. We have proposed⁹ that lesioned animals do not eat because of an insufficient central catecholaminergic response to the severe decrease in glucose utilisation induced by 2-DG (ref. 10). If so, then pretreatments which serve to augment this neurochemical response might be expected to reinstate behavioural function. Consistent with this hypothesis, very large increases in telencephalic tyrosine hydroxylase activity in 6-HDA-lesioned animals, which occur following chronic insulin treatment, are associated with the restoration of 2-DG-induced feeding¹¹. Many of the physiological effects of catecholamines in the sympathetic nervous system seem to be mediated by an increase in the cyclic AMP concentration of the target cells. Methylxanthenes, such as caffeine and theophylline, inhibit phosphodiesterase, prevent cyclic AMP degradation, and thereby potentiate the catecholamine-stimulated rise in cyclic nucleotide^{12,13}. They also enhance many of the behavioural and physiological effects of catecholamines¹²⁻¹⁵, presumably by the same mechanism. We therefore sought to determine whether the acute administration of those sympathomimetic agents, in intact and 6-HDA-lesioned rats, also would potentiate 2-DG-induced feeding, a behaviour that seems to be mediated, in part, by central catecholaminergic neurons. We report that caffeine restores the 2-DG-induced feeding response.

Male albino Sprague-Dawley rats (Zivic-Miller), weighing 250-300 g at the beginning of each experiment, were housed in individual wire mesh cages and allowed free access to

Table 1 Effect of caffeine and 2-DG on food intake

2-DG dose (mg per kg body weight)	n	Food intake above baseline (g)	
		No caffeine	Caffeine
0	6	-0.9±0.5	0.0±0.4
185	5	+2.2±0.9	+5.1±0.5
375	5	+3.2±0.8	+5.2±0.8
750	5	+3.6±0.5	+7.2±1.4

Food intake was measured every hour for 5 h (9.30 a.m. until 2.30 p.m.) on each of two successive days. On day 3, after the first hour of testing, each rat was injected i.p. with isotonic saline or 2-DG. Animals given 2-DG consumed significantly more food (mean±s.e.) during the following 4 h than they had on previous days ($P < 0.01$ in each case). One week later, the rats were tested as before, except that on day 3 they were given caffeine (50 mg per kg, i.p.) along with saline or 2-DG. Rats given 2-DG consumed even more food than they had when 2-DG alone was given ($P < 0.05$ in each case).

Table 2 Effect of xanthenes on 2-DG-induced feeding

Xanthene dose (mg per kg body weight)	n	Food intake (g)	
		Baseline	2-DG + xanthene
None	5	1.3±0.5	3.8±0.7
Caffeine			
12.5	5	0.8±0.3	5.9±0.7
25.0	5	0.9±0.5	6.9±0.5
50.0	5	1.0±0.5	6.3±1.1
100.0	3	0.9±0.5	1.5±0.4
Theophylline			
12.5	5	1.3±0.5	6.9±0.4

Food intake (mean±s.e.) was determined according to the protocol described in Table 1, except that the dose of 2-DG was held constant (375 mg per kg, i.p.) and either the dose of caffeine was varied or theophylline was given instead. All but the highest dose of caffeine increased food intake significantly more than when 2-DG alone was given ($P < 0.01$ in each case).

Purina Lab Chow pellets and tap water. In the first series of experiments animals were given 2-DG, either caffeine or theophylline, or a combination of 2-DG and a xanthene. 2-DG added alone produced a dose-dependent increase in food intake (Table 1). Neither caffeine nor theophylline affected food intake when given alone, but both significantly enhanced the feeding response to 2-DG (Tables 1 and 2).

In the second series of experiments, we administered 6-HDA (200 µg in 20 µl) or its vehicle (0.9% NaCl, 0.1% ascorbic acid) by way of the lateral ventricle 30 min after an intraperitoneal (i.p.) injection of desmethyldipramine (25 mg per kg body weight). We have previously found this treatment to produce a permanent 90-95% depletion of telencephalic dopamine, without altering brain noradrenaline content⁶. Two to three weeks after the lesion was produced, the effects on food intake of optimal doses of 2-DG (375 mg per kg, i.p.) and caffeine (25 mg per kg, i.p.) were determined.

Table 3 Effect of caffeine and 2-DG on food intake in intact and 6-HDA-lesioned rats

	Food intake (g)	
	Intact	6-HDA
Baseline	1.1±0.3	1.1±0.4
2-DG	4.4±1.3	0.7±0.4
2-DG + caffeine	7.7±0.6	7.2±1.0

Food intake (mean±s.e.) was determined according to the protocol described in Table 1, except that a single dose of 2-DG (375 mg per kg, i.p.) was given with caffeine (25 mg per kg, i.p.). Intact rats ($n = 7$) ate more when 2-DG was given with or without caffeine ($P < 0.05$ and < 0.001 , respectively), whereas 6-HDA-lesioned rats ($n = 7$) only ate more when 2-DG was given with caffeine ($P < 0.001$).

Vehicle-treated rats increased food intake in response to 2-DG, and ate even more when caffeine was also administered (Table 3). As expected, 6-HDA-lesioned rats did not increase their food intake after 2-DG treatment alone. When 2-DG was administered together with caffeine, however, the lesioned rats ate significantly more food than before, almost as much as controls did in response to the same treatment (Table 3).

There are several possible explanations for the effects of caffeine on 2-DG-induced feeding in control and lesioned animals. First, caffeine may have altered the effectiveness of 2-DG as a competitive inhibitor by somehow reducing the concentration of circulating glucose. This is apparently not the case, however, since caffeine, if anything, increased the normal hyperglycaemic response to 2-DG (Table 4). Second, because of the ubiquitous involvement of cyclic AMP as a 'second messenger', caffeine may be acting by potentiating the effects of some hormone or transmitter other than brain catecholamines. While the present data do not rule out this possibility, there is a good reason to

believe that 2-DG-induced feeding is mediated in part by the release of central amines since it can be blocked both by the selective destruction of catecholaminergic terminals^{5,6,16} and by AMT (ref. 16), an inhibitor of catecholamine biosynthesis. Third, and most likely, caffeine may produce a net increase in the activation of target cells by catecholamines. This may occur either by potentiating the effects of catecholamines on cyclic AMP levels or by increasing catecholamine release, another known property of methylxanthenes¹⁷. Experiments to further elucidate the mechanisms of this phenomenon are in progress.

Table 4 Effect of caffeine and 2-DG on blood glucose concentration

2-DG dose (mg per kg body weight)	Blood glucose concentration (mg per 100 ml)	
	No caffeine	Caffeine
0	114 ± 5	124 ± 7
185	153 ± 15	201 ± 7
375	274 ± 16	288 ± 14

Blood glucose concentration (mean ± s.e.) in rats after 0 or 25 mg per kg caffeine was given i.p. along with isotonic saline or 2-DG. Food was withheld for 1 h, after which rats were killed by decapitation. From blood samples obtained when rats were killed, glucose was determined colorimetrically following its reaction with glucose oxidase (Worthington) ($n = 4$ or 5 in each group). Statistically significant effects of caffeine were observed only when 185 mg per kg 2-DG had been administered ($P < 0.05$).

Our data are consistent with the hypothesis that 6-HDA-lesioned rats fail to respond behaviourally to 2-DG because of an inadequate activation of central catecholamine receptors. Previously, we also have found that amphetamine will stimulate feeding in anorexic 6-HDA-treated rats⁴, and comparable results have been obtained by others^{18,19}. It is of interest that caffeine and amphetamine both have been used in the treatment of such seemingly diverse syndromes as narcolepsy²⁰ and hyperkinesia^{21,22}. The ability of the sympathomimetic drugs to restore function following damage to central catecholaminergic systems suggests that similar damage may be present in these clinical disorders^{23,24}.

E. M. STRICKER
M. B. ZIMMERMAN
M. I. FRIEDMAN
M. J. ZIGMOND

Departments of Psychology and Life Sciences,
University of Pittsburgh,
Pittsburgh, Pennsylvania 15260

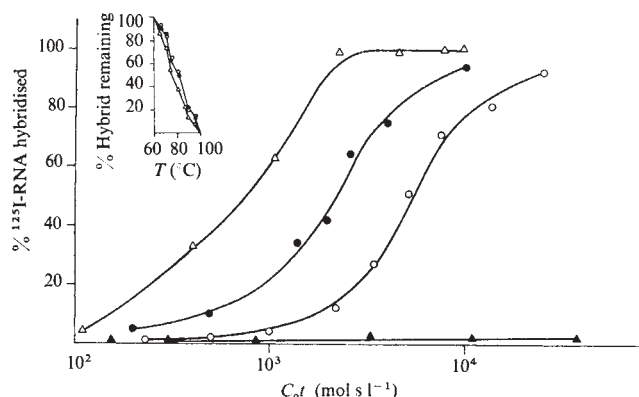
Received 17 January; accepted 8 March 1977.

- Bloom, F. E., Algeri, S., Groppetti, A., Revuelta, A. & Costa, E. *Science* **166**, 1284–1286 (1969).
- Breese, G. R. & Traylor, T. D. *J. Pharmac. exp. Ther.* **174**, 413–420 (1970).
- Uretsky, N. J. & Iversen, L. L. *J. Neurochem.* **17**, 269–278 (1970).
- Stricker, E. M. & Zigmond, M. J. in *Prog. Psychobiol. physiol. Psychol.* **6**, 121–188 (1976).
- Zigmond, M. J. & Stricker, E. M. *Science* **177**, 1211–1214 (1972).
- Stricker, E. M. & Zigmond, M. J. *J. comp. Physiol. Psychol.* **86**, 973–994 (1974).
- Smith, G. P. & Epstein, A. N. *Am. J. Physiol.* **217**, 1083–1087 (1969).
- Stricker, E. M. & Rowland, N. J. *comp. Physiol. Psychol.* (in the press).
- Stricker, E. M., Friedman, M. I. & Zigmond, M. J. *Science* **189**, 895–897 (1975).
- Meldrum, B. S. & Horton, R. W. *Electroenceph. clin. Neurophysiol.* **35**, 59–66 (1973).
- Zigmond, M. J., Volz, C. D., Stricker, E. M. & Friedman, M. I. *Neurosci. Abstr.* **2**, 842 (1976).
- Robison, G. A., Butcher, F. R. W. & Sutherland, E. W. *Cyclic AMP* (Academic, New York, 1971).
- Greengard, P. *Nature* **260**, 101–108 (1976).
- Strömberg, U. & Waldeck, B. *J. Pharm. Pharmac.* **25**, 302–308 (1973).
- Fuxe, K. & Ungerstedt, U. *Med. Biol.* **52**, 48–54 (1974).
- Müller, E. E., Cocchi, D. & Mantegazza, P. *Am. J. Physiol.* **223**, 945–950 (1972).
- Waldeck, B. *J. Pharm. Pharmac.* **23**, 824–830 (1971).
- Teitelbaum, P. & Wolgin, D. L. *Prog. Brain Res.* **42**, 235–249 (1975).
- Ljungberg, T. & Ungerstedt, U. *Physiol. Behav.* **16**, 277–283 (1976).
- Oswald, I. *Pharmac. Rev.* **20**, 273–303 (1968).
- Stroufe, L. A. & Stewart, M. A. *New Engl. J. Med.* **289**, 407–413 (1973).
- Garfinkel, B. D., Webster, C. D. & Sloman, L. *Am. J. Psychiat.* **132**, 723–728 (1975).
- Wender, P. H. *Life Sci.* **14**, 1605–1621 (1974).
- Shaywitz, B. A., Yager, R. D. & Klopfer, J. H. *Science* **191**, 305–308 (1976).

Nucleic acid sequences of primate type C viruses in normal and neoplastic human tissues

RECENT reports on the isolation of type C viruses from human cells have generated considerable controversy because of the close relationship of the viruses to the type C viruses of subhuman primates^{1–5}. These findings have been interpreted as evidence for contamination despite the repeated isolation of some type C viruses from separate clinical specimens or from separate frozen primary cell stocks of the same donor^{6–9}. Our laboratory has described a type C virus which is released from the HEL–12 strain of normal human embryonic lung fibroblasts⁵. The HEL–12 virus is related immunologically to simian sarcoma (woolly monkey fibrosarcoma) virus (SiSV) and the endogenous type C viruses of baboons and domestic cats^{9–11}. To demonstrate that the HEL–12 virus is not an adventitious contaminant, fresh human tissues were examined for antigenic cross-reactivity with HEL–12 virus antigens. Some patients with myelogenous leukemia contain circulating immunoglobulins which specifically inhibit the reverse transcriptases of HEL–12 virus and primate type C viruses¹². In addition, the glomerular immune complexes in patients with systemic lupus erythematosus contain antigens related to HEL–12 virus proteins¹³. Using the techniques of molecular hybridisation, we demonstrate here that HEL–12 cells contain proviral DNA sequences before antigen expression or spontaneous virus release can be detected. We also present evidence for the existence of nucleic acid sequences homologous to HEL–12 viral RNA in the DNAs from certain cancer patients.

HEL–12 cultures were grown from frozen primary cell stocks⁵. Type C viral antigens or extracellular virus were not detected during the first 20–40 d, but the cultures gradually expressed cytoplasmic viral antigens and released type C particles between 80 and 120 d. DNA was extracted from early-passage (36 d of culture) HEL–12 cells in which type C viral antigens were not detected by indirect immunofluorescence or from late-passage (115 d of culture) HEL–12 cells which spontaneously released type C virus^{9,14}. The DNA was sheared, denatured and reannealed in the presence of trace amounts ¹²⁵I-HEL–12 viral RNA. Aliquots were evaluated for duplex formation at various times thereafter by observing resistance to pancreatic RNase digestion in conditions of high ionic strength. Both early-passage and late-passage HEL–12 cells contained DNA sequences homologous to HEL–12 viral RNA (Fig. 1). The relative displacement of the curves along the abscissa suggested that early-passage cells contained fewer copies of the viral genome than late-passage cells. The ¹²⁵I-HEL–12 RNA also hybridised to DNA from a marmoset fibroblast line infected with SiSV (71API/SiSV), thereby confirming the relatedness of HEL–12 virus to SiSV⁵. The ¹²⁵I-HEL–12 viral RNA probe eventually hybridised to a similar extent with DNA from early-passage and late-passage HEL–12 cells and 71API/SiSV cells. Some differences were noted between the t_m values of the hybrids formed with HEL–12 viral DNAs and 71API/SiSV-DNA, suggesting that the proviral sequences being detected were not identical in all cases (Fig. 1, Table 1). Since HEL–12 cell cultures express antigens related to SiSV and baboon endogenous virus (BaEV)^{9,14}, HEL–12 viral RNA was hybridised to DNA extracted from a bat lung cell line productively infected with BaEV. The stock of HEL–12 viral RNA used for these experiments lacked detectable BaEV sequences (Table 1), possibly because the HEL–12 cell cultures released virions with SiSV in great excess. Moreover, we have observed intervals in the lifespan of HEL–12 cells during



(Grade T-2, Collaborative Research) was added. The mixture was stirred at room temperature for 16 h. The oligo(dT)-cellulose was removed and packed into a Pasteur pipette which was washed exhaustively with $2\times$ SSC. The poly(A)-containing HEL-12 viral RNA was eluted by successive washes of the column with 10 mM Tris-HCl, pH 7.2. The eluate was adjusted to a final NaCl concentration of 0.25 M in order to stabilise the RNA. Iodination was performed as originally described by Commerford¹⁸ and modified by Prenskey *et al.*¹⁰. The specific activity of the ^{125}I -RNA was approximately $5\text{--}10\times 10^7$ c.p.m. per μg . Hybridisations were carried out in sealed glass capillary tubes with approximately 1,000 c.p.m. of ^{125}I -viral RNA in 0.4 M Na-phosphate, pH 6.8. The DNA-RNA ratio in all cases was at least 5×10^6 . Each reaction mixture was placed in boiling water for 10 min and subsequently incubated at 60 °C. Hybrids were evaluated at progressively greater C_0t values²⁰ by determining resistance to $20\text{ }\mu\text{g ml}^{-1}$ pancreatic RNase (Worthington) at high ionic strength ($2\times$ SSC) for 2 h at 37 °C. Less than 3% of the ^{125}I -viral RNA was resistant to RNase when incubated in the absence of DNA or when incubated to $C_0t = 4\times 10^4$ with calf thymus DNA. This resistant fraction was subtracted from all values. The C_0t values shown are uncorrected for $[\text{Na}^+]$ ion concentration. The final extent of hybridisation is plotted relative to an absolute value of 43% obtained using 71AP1/SiSV DNA. Inset, thermal stabilities of ^{125}I -RNA:DNA hybrids. ^{125}I -HEL-12 viral RNA and DNAs from various sources were annealed to a C_0t of $10^4\text{--}2\times 10^4$ and diluted into 50–100 volumes of $1\times$ SSC buffer. Aliquots were held at the indicated temperatures for 10 min, adjusted to a final concentration of $2\times$ SSC buffer and digested with $20\text{ }\mu\text{g ml}^{-1}$ pancreatic RNase for 2 h at 37 °C. Sources of DNAs: Δ , marmoset fibroblast cells releasing SiSV (71AP1/SiSV); $\circ$, HEL-12 cells passed for 36 d; $\bullet$, HEL-12 cells passed for 115 d; $\blacktriangle$, calf thymus DNA.

which only SiSV-related antigens are expressed¹⁴. The absence of BaEV-RNA in this stock of HEL-12 virus has been confirmed by R. G. Smith and R. C. Gallo (personal communication).

Two of 14 DNA samples from nine cancer patients likewise contained DNA sequences hybridisable to the RNA of HEL-12 virus (Fig. 2). In one case, proviral sequences were detected in the spleen of an 11-yr old boy who died from an osteogenic sarcoma. This tumour uncommonly metastasises to the spleen²¹, and no evidence for splenic metastases was obtained by histological post-mortem examination. The results with DNAs from other tissues of this patient were interesting because HEL-12 proviral sequences were not detected in the primary tumour from the femur

or in the patient's liver (Fig. 2). A second case involved the spleen of an adult with acute myelogenous leukaemia. Proviral DNA sequences were found in the spleen although to a lesser extent than in early-passage or late-passage HEL-12 cells or in the spleen of the boy with osteogenic sarcoma (Fig. 2). The thermal stabilities of these hybrids were not appreciably different from those of the ^{125}I -RNA:HEL-12 DNA hybrids (Fig. 2, Table 1).

The results (Table 1) can be summarised as follows: (1) DNA extracted from early-passage HEL-12 cells contain proviral sequences. (2) With continued *in vitro* propagation, HEL-12 cells acquire an increased complement of proviral sequences, possibly as the result of a selective proliferation of a cell fraction containing the HEL-12 viral genome or

Table 1 HEL-12 proviral sequences in various DNAs

Diagnosis	Source of DNA	Final relative % hybridisation of ^{125}I -HEL-12 viral RNA	Hybrid t_m (°C)	Δt_m (°C)
—	71AP1/SiSV	100	77	8.0
—	HEL-12 cells, early-passage	95	82.5	2.5
—	HEL-12 cells, late-passage	93	82.0	3.0
Osteogenic Sarcoma	Spleen	90	84.5	0.5
	Liver	1	—	—
	Primary tumour	1	—	—
AML	Spleen	61	85.0	0
	Liver	2	—	—
AML	Spleen	4	—	—
AML	Liver	3	—	—
AML	Kidney	2	—	—
	Liver	2	—	—
AML	Peripheral leukocytes	1	—	—
AML	Peripheral leukocytes	3	—	—
ALL	Kidney	2	—	—
Nasopharyngeal Carcinoma	Liver	4	—	—
	Spleen	0	—	—
—	BL88/BaEV	6	—	—

The actual % hybridisation of 71AP1/SiSV DNA was 43. The Δt_m (°C) is expressed relative to a t_m of 85.0 °C for ^{125}I -HEL-12 viral RNA:AML spleen DNA hybrid. AML, acute myelogenous leukaemia; ALL, acute lymphatic leukaemia; BL88/BaEV, bat lung cells infected with BaEV from *Papio anubis*. All DNAs were annealed with ^{125}I -HEL-12 viral RNA to a $C_0t \geq 10^4$.

from a low level of virus spread from cell to cell in early-passage cultures. Embryonic tissues were not available to demonstrate proviral sequences in the original lung tissue from which the HEL-12 cell strain was established in 1974. Even if available, proviral sequences might not be found if present in a small number of cells of such heterogeneous tissues. (3) DNA sequences homologous to HEL-12 viral RNA have also been detected in the spleen DNA of a boy with osteogenic sarcoma, and to a lesser extent, in the spleen of an adult with acute myelogenous leukaemia. The proviral sequences in the DNAs of HEL-12 cells and fresh human tissues were related to SiSV as evidenced by the ability of the ^{125}I -HEL-12 viral RNA to hybridise to 71AP1/SiSV DNA. Thermal denaturation profiles of these hybrids, however, reveal some differences between the hybrids formed between HEL-12 viral RNA and DNA from cells producing HEL-12 virus or SiSV. This suggests

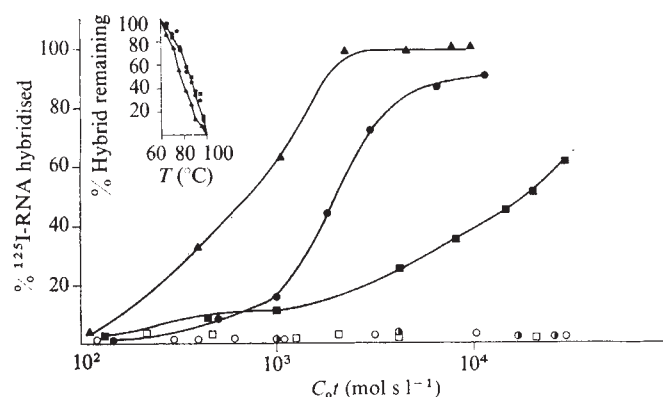


Fig. 2 Hybridisation of ^{125}I -HEL-12 viral RNA to DNAs from organs of cancer patients. The conditions for purification and shearing of DNAs were as described in Fig. 1. The final extent of hybridisation is plotted relative to that obtained with 71AP1/SiSV DNA. Source of DNAs: Δ , 71AP1/SiSV (from Fig. 1); $\bullet$, spleen, and $\circ$, liver from patient with osteogenic sarcoma; $\blacksquare$, primary osteogenic sarcoma; $\blacksquare$, spleen and $\square$, liver from patient with acute myelogenous leukaemia. Inset, thermal stabilities of hybrids in $1\times\text{SSC}$ buffer.

that the proviral sequences in HEL-12 cells and fresh human tissues may not be identical to those detected in 71AP1/SiSV cells. The presence of proviral sequences in early-passage HEL-12 cells argues against the notion that the HEL-12 virus is a laboratory contaminant, especially since primate type C viruses have not been propagated by us⁵. But, by far the strongest argument against adventitious contamination seems to be the presence of SiSV-related sequences in fresh autopsy tissues of patients with malignancies.

We thank R. C. Gallo, D. H. Gillespie, S. Panem and R. G. Smith for helpful discussions, W. Prenskey for a gift of iodinated RNA and M. Reitz for DNA of 71AP1/SiSV. This work was supported by the US NCI. E. V. P. is supported by the National Institute of General Medical Sciences.

EDWARD V. PROCHOWNIK
WERNER H. KIRSTEN

Departments of Pathology and Pediatrics,
University of Chicago,
Chicago, Illinois 60637

Received 16 December 1976; accepted 14 March 1977.

- 1 Mak, T. W., Kurtz, S., Manaster, J. & Housman, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 623 (1975).
- 2 Gallagher, R. E. & Gallo, R. C. *Science* **187**, 350 (1975).
- 3 Nooter, K., Aarsen, A. M., Bentvelzen, P., DeGroot, F. G. & Van Pelt, F. G. *Nature* **256**, 595 (1975).
- 4 Gableman, N., Waxman, S., Smith, W. & Douglas, S. *Int'l J. Cancer* **16**, 355 (1975).
- 5 Panem, S., Prochownik, E. V., Reale, F. & Kirsten, W. H. *Science* **189**, 297 (1975).
- 6 Teich, N. M. *et al. Nature* **256**, 551 (1975).

- 7 Reitz, M. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2113 (1976).
- 8 Wong-Staal, F., Gillespie, D. H. & Gallo, R. C. *Nature* **262**, 190 (1976).
- 9 Panem, S., Prochownik, E. V. & Kirsten, W. H. in *Comparative Leukemia Research 1975* (ed. Clemmesen, J. & Yohn, D. S.) 572 (Karger, Basel, 1976).
- 10 Theilen, G. H., Gould, D., Fowler, M. & Dungworth, D. *J. natn. Cancer Inst.* **47**, 881 (1971).
- 11 Sherr, C. J. & Todaro, G. J. *Virology* **61**, 168 (1974).
- 12 Prochownik, E. V. & Kirsten, W. H. *Nature* **260**, 64 (1976).
- 13 Panem, S., Ordóñez, N., Kirsten, W. H., Katz, A. I. & Spargo, B. H. *New Engl. J. Med.* **295**, 470 (1976).
- 14 Panem, S., Prochownik, E. V., Knish, W. & Kirsten, W. H. *J. gen. Virol.* (in the press).
- 15 Smith, R. G., Donehower, L., Gallo, R. C. & Gillespie, D. H. *J. Virol.* **17**, 287 (1976).
- 16 Schwartz, S. & Kirsten, W. H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3570 (1974).
- 17 Tibbets, C., Johansson, K. & Philipson, L. *J. Virol.* **12**, 218 (1973).
- 18 Commerford, S. L. *Biochemistry* **10**, 1993 (1971).
- 19 Prenskey, W., Steffensen, D. M. & Hughes, W. L. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1860 (1973).
- 20 Britten, R. J. & Kohne, D. E. *Science* **161**, 529 (1968).
- 21 Evans, R. W. *Histological Appearances of Tumours*, 2nd edn (Livingstone, London, 1966).

Chaos in an enzyme reaction

DYNAMIC systems are usually thought to have either monotonic or periodic behaviour. Although the possibility of other types of behaviour has been recognised for many years, the existence of non-monotonic, non-periodic behaviour in dynamic systems has been firmly established only recently. It is termed chaotic behaviour. A review on the rapidly expanding literature on chaos in discrete model systems described by difference equations has been published by May¹. Rössler², on the other hand, has discussed a few published works on systems of differential equations with chaotic solutions, and he has proposed a three-component chemical model system which he argues has chaotic

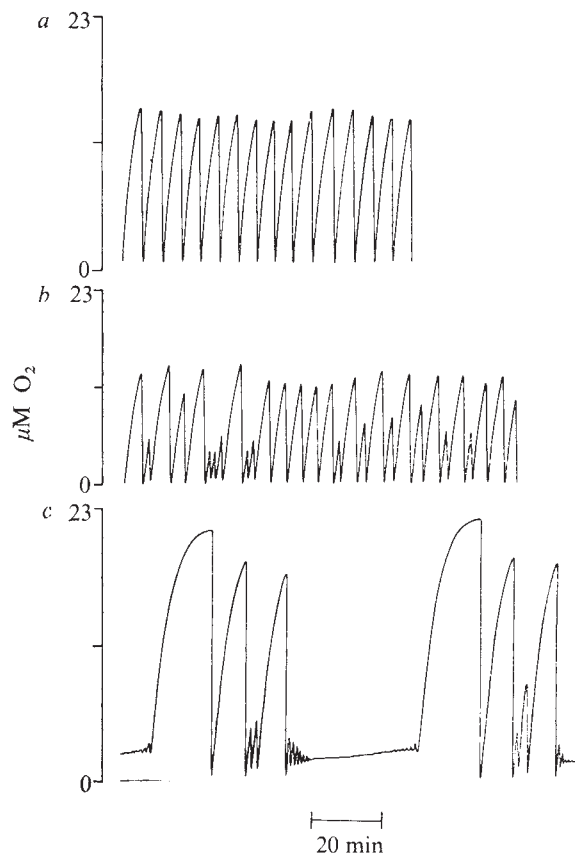


Fig. 1 Oscillations in oxygen concentration in peroxidase catalysed oxidation of NADH in system open to O_2 . A stirred 5 ml sample containing horseradish peroxidase, 10 μM 2,4-dichlorophenol and 0.2 μM methylene blue in a 0.1 M Na-acetate buffer pH 5.1 was in contact with an O_2/N_2 gas mixture at atmospheric pressure containing 1.9 volume % O_2 . A solution of 0.2 M NADH was pumped into the sample at 11 μl per h. The oxygen concentration in the liquid was measured with a Clark electrode. The temperature was 28 $^\circ\text{C}$. The concentration of peroxidase was a, 0.90 μM ; b, 0.55 μM and c, 0.45 μM .

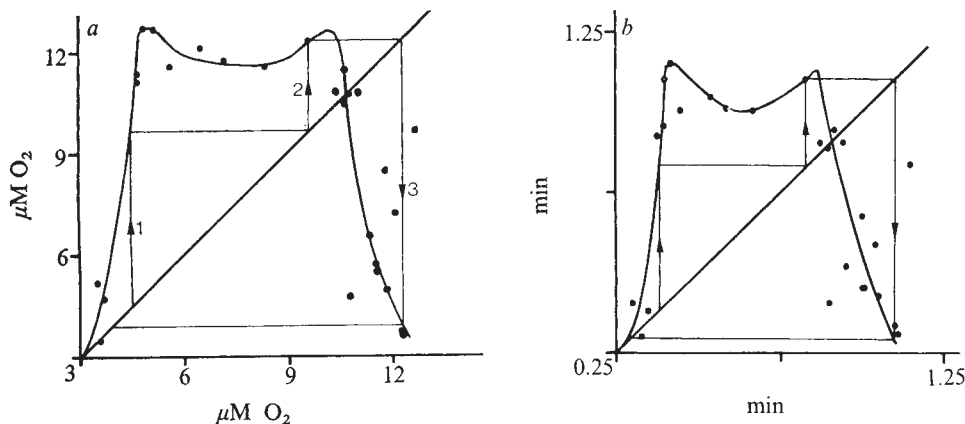


Fig. 2 *a*, Amplitudes ($A(n+1)$) plotted against preceding amplitudes ($A(n)$) from Fig. 1*b*. *b*, Periods ($P(n+1)$) plotted against preceding periods ($P(n)$) from Fig. 1*b*. Trajectories with arrows were drawn to show that the transition functions allow the period 3.

solutions. The argument is based on a theorem by Li and Yorke³. Here we report the finding of chaotic behaviour as an experimental result in an enzyme system (peroxidase). Like Rössler² we base our identification of chaos on the theorem by Li and Yorke³.

When the oxidation of NADH by O_2 , catalysed by horseradish peroxidase, takes place in a stirred solution where O_2 can enter by diffusion from the gas phase, oscillations^{4,5} and bistability⁶ may be observed in the reaction system. Sustained oscillations may be achieved by continuously supplying NADH to the reaction mixture. In the present experiments this was done by pumping in an NADH solution at a slow constant rate. The wave form of the oscillation depended on the enzyme concentration as shown in Fig. 1. In Fig. 1*b* there is no apparent periodicity. This irregular oscillation was analysed by plotting each amplitude against the preceding amplitude as described by Lorenz⁷. This plot (Fig. 2*a*) reveals that the amplitudes found in our experiment satisfy a well defined cap-shaped transition function. When the durations of the periods are plotted in the same way a similar transition function is found (Fig. 2*b*). A random distribution of points in the plane would have been expected if the irregularities in the oscillation were caused by random perturbations.

The existence of the cap-shaped transition function indicates that the system has either a periodic behaviour with a complicated period or it has chaotic behaviour. The two possibilities can be distinguished using the theorem of Li and Yorke³ which states that if the transition function allows the period three it allows any period and chaos exists. In Fig. 2*a* and *b* it is seen that it is possible to design trajectories of three transitions which will bring the system back beyond the origin. Thus our system exhibits chaotic behaviour in the Li and Yorke sense. We are not aware of any other examples of chaos in chemical reaction systems.

LARS FOLKE OLSEN
HANS DEGN

Institute of Biochemistry,
Odense University,
Odense, Denmark

Received 1 November 1976; accepted 14 March 1977.

- ¹ May, R. M. *Nature* 261, 459–467 (1976).
- ² Rössler, O. E. *Z. Naturforsch.* 31a, 259–264 (1976).
- ³ Li, T.-Y. & Yorke, J. A. *Am. math. Mon.* 82, 985–992 (1975).
- ⁴ Nakamura, S., Yokota, K. & Yamazaki, I. *Nature* 222, 794 (1969).
- ⁵ Degn, H. *Biochim. biophys. Acta* 180, 271–290 (1969).
- ⁶ Degn, H. *Nature* 217, 1047–1050 (1968).
- ⁷ Lorenz, E. N. *J. atmos. Sci.* 20, 130–141 (1963).

Specific inhibition of chromatin-associated poly(A) synthesis *in vitro* by cordycepin 5'-triphosphate

It has been established that many eukaryotic mRNAs contain poly(adenylic acid) tracts at their 3'-termini. The polyadenylation of mRNA occurs post-transcriptionally in the nucleus as a rapid, initial addition of 100–200 adenylate residues to the pre-mRNA (ref. 1). Subsequently, a slower chain extension (6–8 bases) of the poly(A) tail seems to occur both in the nucleus and in the cytoplasm². The initial polyadenylation reaction can be specifically inhibited by the drug cordycepin (3'-deoxyadenosine) in cell culture, presumably by its conversion to the triphosphate analogue which acts as a competitive inhibitor of poly(A) polymerase. Cordycepin, however, has little effect on the slower poly(A) extension reaction³ or on the formation of mRNA precursor molecules^{4–6}; but it can inhibit rRNA synthesis⁷. Contrary to the *in vitro* observations, cordycepin 5'-triphosphate (3'dATP) is not a specific inhibitor of poly(A) synthesis *in vivo*, relative to RNA synthesis^{8,9}, and RNA polymerase I (which synthesises rRNA) is actually less sensitive to inhibition by 3'dATP than RNA polymerase II (ref. 10) (which is presumed to be involved in the synthesis of mRNA). Since nuclear poly(A) polymerase occurs in two functional states as 'free' and 'chromatin-bound' forms¹¹, we reasoned that if the chromatin-associated poly(A) polymerase were involved in the initial polyadenylation of mRNA, it might be selectively inhibited by 3'dATP. The present studies, designed to test such an idea, demonstrate that, as *in vivo*, the initial polyadenylation reaction can be selectively inhibited *in vitro* by low levels of 3'dATP. These data also show that higher levels of 3'dATP can inhibit RNA synthesis, 'chromatin-bound' RNA polymerase I activity being significantly more sensitive than the 'bound' RNA polymerase II activity.

Initial studies measuring the effect of 3'dATP on poly(A) synthesis catalysed by 'bound' and 'free' poly(A) polymerases in whole nuclei, indicated that the 'bound' enzyme was 100–200 times more sensitive than the 'free' poly(A) polymerase. To study the inhibition by 3'dATP in more detail, we separated the poly(A)-synthesising activities of isolated nuclei by preparing a nuclear sap fraction and chromatin from the nuclear residue to yield the 'free' and 'bound' enzymes, respectively. The effect of increasing concentrations of 3'dATP on poly(A) synthesis catalysed by the 'free' and by the 'bound' enzyme is shown in Fig. 1*a*. Approximately 150 $\mu\text{g ml}^{-1}$ of drug were needed to achieve 50% inhibition of the 'free' enzyme, whereas the same inhibition was attained with less than 1 $\mu\text{g ml}^{-1}$ of 3'dATP

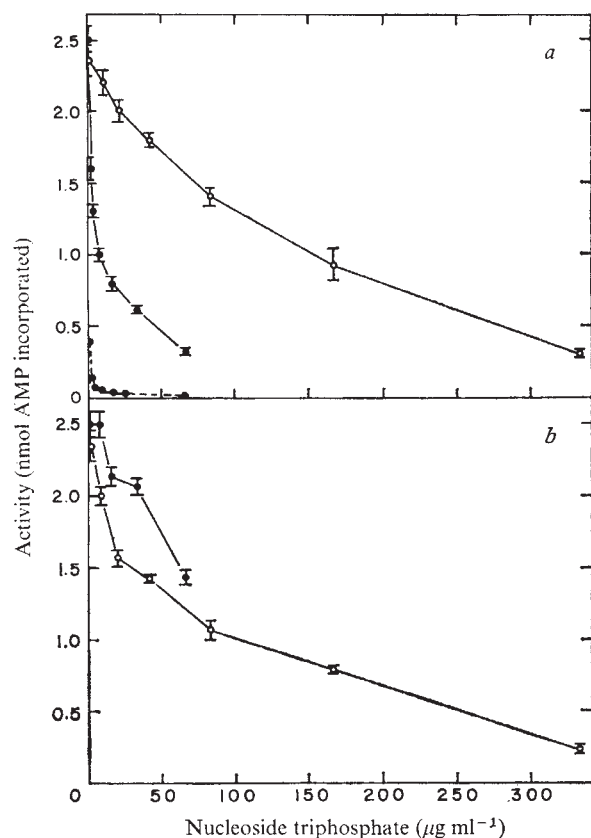


Fig. 1 Effect of 3'dATP (a) and 2'dATP (b) on 'free' and 'bound' poly(A) polymerases. Rat liver nuclei were prepared by homogenisation of the tissue in hypertonic sucrose (2.3 M) (ref. 15) and the isolated nuclei were gently resuspended in hypotonic buffer essentially as described previously¹¹. After centrifugation (40,000g; 30 min) to remove the nuclei, glycerol was added to the supernatant fraction so that the final concentration was 20% (v/v). This nuclear supernatant fraction was the source of the 'free' poly(A) polymerase. Chromatin was prepared from the nuclear pellet after washing with the same hypotonic buffer. Isolation of chromatin was as described previously¹⁶ except that the Triton X-100 and NaCl washes were omitted. The final chromatin pellet was resuspended in 0.4 mM Tris-HCl (pH 8) (0.5 ml per g liver) and used within 36 h of preparation. The DNA content of nuclei and chromatin was estimated as described by Burton¹⁷. Poly(A) polymerase activity was assayed in a reaction mixture containing 38 mM Tris-HCl (pH 8), 38 mM KCl, 0.7 mM MnCl₂, 1.0 mg per ml phosphoenolpyruvate, 2 μg per ml pyruvate kinase, 0.37 mM ³H-ATP (2,000 c.p.m. nmol⁻¹), addition of 3'dATP or 2'dATP as indicated and appropriate amounts of poly(A) polymerase in a total volume of 120 μl. For analysis of 'free' poly(A) polymerase, poly(A) was included at a concentration of 375 μg ml⁻¹. After 1 h at 37 °C, the reaction was terminated by addition of excess ATP (10 μl of ATP, 10 mg ml⁻¹). Aliquots of 110 μl were spotted on to DE81 filters and processed as described previously¹⁵. All assays were performed in quadruplicate; zero time backgrounds were less than 150 c.p.m. ○, 'Free' poly(A) polymerase obtained from nuclei representing 45 μg DNA; ●, chromatin representing 50 μg (---) or 160 μg (—) DNA. Horizontal bars indicate s.e.m.

when 'bound' poly(A) polymerase was derived from an equivalent amount of tissue. Since the activity catalysed by the chromatin-associated enzyme represents only a small percentage of the activity of the 'free' poly(A) polymerase, we also measured the effect of 3'dATP when the levels of chromatin and nuclear supernatant (sap) fractions were adjusted to give the same activity. When the activities were so adjusted, the 'bound' enzyme was still 75 times more sensitive than the 'free' enzyme.

In order to further substantiate the specific inhibitory effect of 3'dATP on chromatin-associated poly(A) polymerase relative to 'free' enzyme, the effect of 2'dATP, the naturally occurring deoxynucleotide, was measured. 2'dATP

did not specifically inhibit chromatin-bound poly(A) synthesis (Fig. 1b); in fact, 'free' poly(A) polymerase was slightly more sensitive to inhibition by 2'dATP than the 'bound' enzyme. Interestingly, inhibition of 'free' poly(A) polymerase by 2'dATP was greater than that obtained with 3'dATP, 50% inhibition occurring at 70 and 120 μg ml⁻¹, respectively.

Since 3'dATP is an ATP analogue, the relationship of the inhibitor to substrate levels was investigated. Figure 2 shows the double reciprocal plots of velocity against ATP concentration in the presence and absence of 3'dATP. The inhibition of the 'free' poly(A) polymerase by 3'dATP was competitive with substrate (b). The K_m and K_i calculated from Fig. 2b were 62 μM and 40 μM, respectively. When the relationship of 3'dATP to ATP levels was investigated using the chromatin fraction, a non-competitive relationship was observed (Fig. 2a). The calculated K_m and K_i were 100 μM and 1.3 μM, respectively. Although the K_m values of the two enzyme activities were similar, the K_i for chromatin-associated poly(A) synthesis was much lower, indicating that this reaction was at least 30 times more sensitive to inhibition by 3'dATP than poly(A) synthesis by the 'free' enzyme. Inhibition of poly(A) polymerase for both the 'free'

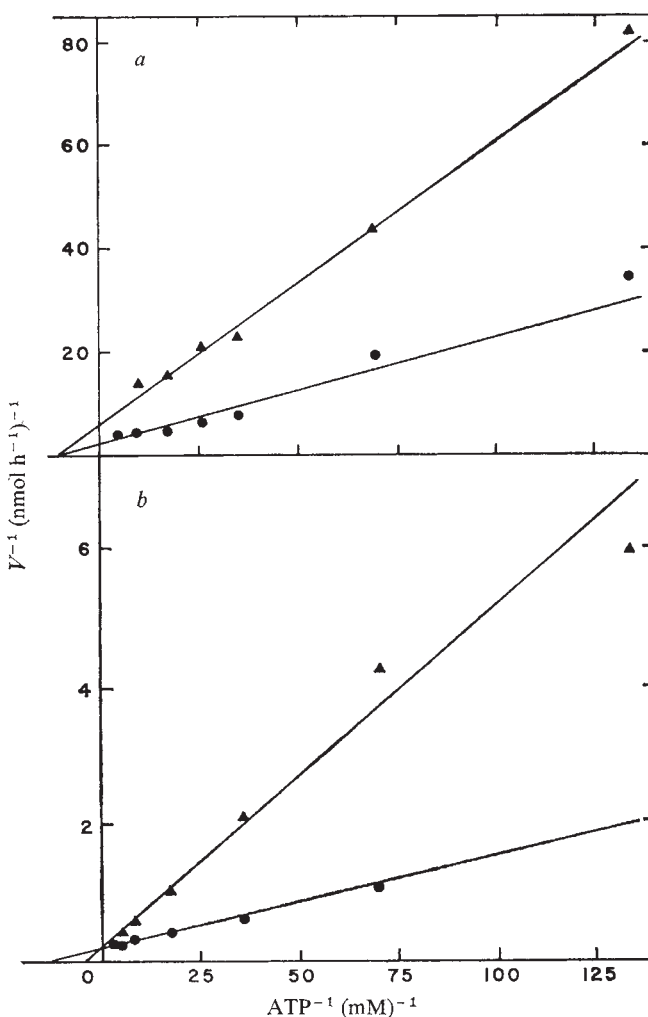


Fig. 2 Effect of 3'dATP on the double reciprocal plot of velocity against substrate concentration for 'bound' and 'free' poly(A) polymerases. Chromatin and nuclear supernatant fractions were prepared and assayed as described in Fig. 1 in the presence (▲) and absence (●) of 3'dATP at varying ³H-ATP⁻¹ levels. a, Chromatin equivalent to 60 μg DNA and 3'dATP at a concentration of 4 μM were used; b, 'free' poly(A) polymerase was derived from 90 μg nuclear DNA and 3'dATP was at a concentration of 118 μM. The ATP concentration in the assay ranged from 7–235 μM with specific activities ranging from 2 × 10⁵ to 7 × 10⁵ c.p.m. nmol⁻¹. Lines were fitted by least squares analysis.

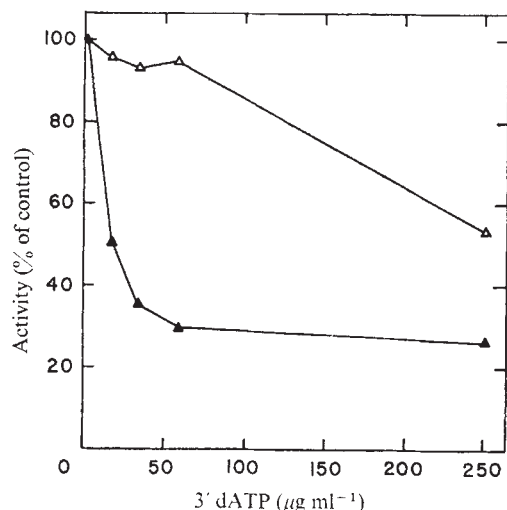


Fig. 3 Effect of 3'dATP on chromatin-bound RNA polymerases. RNA synthesis was measured in an assay similar to that described in Fig. 1 except that CTP, GTP and ^3H -UTP (30 c.p.m. pmol $^{-1}$) were included at a concentration of 50 μM . Chromatin equivalent to 60 μg DNA was used. Reactions were incubated for 5 min at 30 °C, after which excess UTP was added and filters processed as described previously¹⁵. Reactions were performed in quadruplicate; zero times backgrounds were less than 100 c.p.m. Analyses were performed in the presence and absence of α -amanitin. The decrease in activity when 1 $\mu\text{g ml}^{-1}$ α -amanitin was included represented RNA polymerase II. Activity in the presence of α -amanitin represented primarily RNA polymerase I since high levels of toxin (150 $\mu\text{g ml}^{-1}$) produced only 10–15% decrease in activity (indicating little RNA polymerase III contamination). Δ , RNA polymerase II, 100% activity corresponded to 22.8 ± 1.8 pmol UMP incorporated in 5 min; $\blacktriangle$, RNA polymerase I, 100% activity corresponded to 13.2 ± 0.8 pmol UMP incorporated in 5 min.

and 'bound' forms by 2'dATP was competitive with ATP levels, so it seems that not only is 3'dATP a specific inhibitor of chromatin-associated poly(A) synthesis, but that its mode of action differs from inhibition of the 'free' form of enzyme.

We then measured the effect of 3'dATP on RNA polymerase activity associated with the chromatin fraction, maintaining the assay conditions as close as possible to those for poly(A) synthesis to avoid unnecessary artefacts. Figure 3 shows the inhibition of chromatin-associated RNA polymerases I and II by 3'dATP. It can be seen that polymerase I was more sensitive to the inhibitor than polymerase II. At 60 $\mu\text{g ml}^{-1}$ 3'dATP, which inhibited polymerase I activity by 70%, polymerase II activity was inhibited only by 5%. Although the 'bound' polymerase I activity was reduced by 3'dATP, this inhibition was far less than observed for 'bound' poly(A) polymerase activity, where 50% inhibition could be achieved at concentrations of less than 1 $\mu\text{g ml}^{-1}$ using a similar amount of chromatin. RNA polymerase activity can also be detected in the nuclear supernatant fraction¹². As previously reported for partially purified enzymes^{8,10}, 3'dATP had a similar effect on 'free' RNA polymerases present in the nuclear supernatant fraction, with a dose-response curve similar to that for chromatin-bound RNA polymerase I (data not shown).

The marked sensitivity of chromatin-associated poly(A) polymerase, the lack of inhibition of RNA polymerase II in the same subnuclear fraction and the intermediate sensitivity of chromatin-bound RNA polymerase I by 3'dATP are results similar to those obtained *in vivo* by cordycepin. The lack of differential inhibition of 'free' poly(A) polymerase relative to 'free' RNA polymerases by 3'dATP supports the observations made using partially purified enzymes. Thus, although purified enzymes are invaluable in determining reaction characteristics, data obtained from such systems

should be interpreted with caution. In this case, inhibition of partially purified enzymes by 3'dATP, probably reflects the effects of the drug on extrachromosomal events. The selectivity of inhibition of chromatin-associated poly(A) synthesis by 3'dATP is further evident when a comparison is made with the effect of the compound on polyadenylation of endogenous RNA in isolated mitochondria, where chromatin factors are not present. Inhibition of the mitochondrial polyadenylation reaction occurs only at relatively high concentrations of 3'dATP (ref. 13) (50% inhibition occurring at 300 μM , a value even greater than the K_m). The lack of inhibition of mitochondrial polyadenylation at this concentration of 3'dATP agrees well with the relative insensitivity of mitochondrial polyadenylation to cordycepin *in vivo*¹⁸.

The remarkable sensitivity of chromatin-associated poly(A) polymerase to cordycepin triphosphate, the close association of this activity with DNA and the similar inhibition responses of the initial polyadenylation reaction *in vivo* and 'bound' enzyme activity *in vitro*, all tend to indicate that the latter activity indeed represents the initial polyadenylation reaction. The 'free' poly(A) polymerase activity, by virtue of its localisation in the nuclear supernatant and cytosol fractions¹⁴, coupled with its relative insensitivity to low levels of 3'dATP, should represent the enzyme activity responsible for the additional polyadenylation reaction. These studies have thus provided a simple system to differentiate between initial polyadenylation and subsequent chain extension reactions *in vitro*.

This work was supported by research grants from United States Public Health Service.

KATHLEEN M. ROSE
LINN E. BELL
SAMSON T. JACOB

Department of Pharmacology,
Milton S. Hershey Medical Centre,
Pennsylvania State University,
College of Medicine,
Hershey, Pennsylvania 17033

Received 27 January; accepted 27 March 1977.

- Darnell, J. E., Jelinek, W. R. & Molloy, G. R. *Science* **181**, 1215–1221 (1973).
- Brawerman, G. *Prog. Nucl. Acid. Res. molec. Biol.* **17**, 117–148 (1976).
- Diez, J. & Brawerman, G. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4091–4095 (1974).
- Penman, S., Rosbash, M. & Penman, M. *Proc. natn. Acad. Sci. U.S.A.* **67**, 1878–1885 (1970).
- Darnell, J. E., Philipson, L., Wall, R. & Adesnik, M. *Science* **174**, 507–510 (1971).
- Mendecki, J., Lee, S. & Brawerman, G. *Biochemistry* **11**, 792–798 (1972).
- Siev, M., Weinberg, R. & Penman, S. *J. Cell Biol.* **41**, 510–519 (1969).
- Horowitz, B., Goldfinger, B. & Marmur, J. *Archs Biochem. Biophys.* **172**, 143–148 (1976).
- Maale, G., Stein, G. & Mans, R. *Nature* **255**, 80–82 (1975).
- Desrosiers, R., Rottman, F., Boezi, J. & Towle, H. *Nucl. Acids Res.* **3**, 325–341 (1976).
- Jacob, S. T., Roe, F. J. & Rose, K. M. *Biochem. J.* **153**, 733–735 (1976).
- Yu, F.-L. *Biochim. biophys. Acta* **395**, 329–336 (1975).
- Rose, K. M. & Jacob, S. T. *Biochemistry* **15**, 5046–5052 (1976).
- Rose, K. M., Lin, Y.-Ch. & Jacob, S. T. *FEBS Lett.* **67**, 193–197 (1976).
- Rose, K. M., Ruch, P. A., Morris, H. P. & Jacob, S. T. *Biochim. biophys. Acta* **432**, 60–72 (1976).
- Jänne, O., Bardin, C. W. & Jacob, S. T. *Biochemistry* **14**, 3589–3597 (1975).
- Burton, K. *Biochem. J.* **62**, 315–323 (1956).
- Hirsch, M. & Penman, S. *J. molec. Biol.* **83**, 131–142 (1974).

Nature Index and Binders

The complete **Index** for 1976 is available, price £2.50. Copies of the 1975 index are still on sale, price £3.00. **Binders** for the journal are also available at £8.00 for four (a year of *Nature* fits into four binders).

Postage is included in the above prices. Orders should be sent, accompanied by remittance to Macmillan Journals Ltd, Brunel Road, Basingstoke, Hampshire, England.

matters arising

The Late Weichselian geomagnetic event

THOMPSON and Berglund¹ describe palaeomagnetic results from Björkeröds Mosse which, they claim, invalidate all previous evidence for a Late Weichselian geomagnetic event. In this note I discuss these results and answer their criticisms in regard to evidence which I have presented².

Thompson and Berglund stress the difficulty of interpreting palaeomagnetic records from sedimentologically complex cores and yet the 'carefully selected site' at Björkeröds Mosse contains eight sedimentary units from three climatic zones. The high noise content of the magnetisation, evident in Fig. 1 of ref. 1, thus supports their contention that variable sediments are poor recorders of the ambient field and this is confirmed by the absence of any secular variation periodicity comparable with cycles in archaeomagnetic data³. Their scattered directions have probably arisen from the spasmodic growth of post-depositional magnetisation, while the general decrease in inclination with depth through layers 8 to 2 is consistent with increasing compaction in these organic sediments^{4,5}. These factors question the reliability of the Björkeröds Mosse data and thus the confidence with which the authors reject all previous results.

In contrast, I have described evidence for a Late Weichselian excursion from a single lithology of exceptionally uniform clay varves from Stjärnö, Sweden². Studies of similar varves^{5,6} show that such clays were magnetised depositionally and if sedimentation occurred on a horizontal surface in the absence of strong water currents then the remanence is an accurate record of the ambient field. This is shown, for example at Stjärnö, by the virtual absence of noise in varves which satisfy these criteria. The Revised Swedish Time Scale also provides a precise chronology within which to assemble the history of an excursion. At Stjärnö the varve data were characterised by a reversal of declination and a shallowing, but not a reversal, of inclination while elsewhere in Sweden southerly declinations with

positive inclinations are a feature of younger varved clays (~8,000–7,600 BC) (ref. 5). As the Björkeröds Mosse core was unoriented, declination values are arbitrary and cannot be used to contradict these observations. Furthermore, according to Berglund⁷ it is impossible to correlate the varved clays in layer 1 with the Swedish Time Scale thus raising doubts concerning any direct age comparison between the Björkeröds Mosse data and results from dated varves.

Thompson and Berglund are correct in emphasising the need for caution when interpreting palaeomagnetic information from sediments, but their suggestion of using the repeatability of between-core directions as a major criterion for confirming the direction of the palaeofield could be misleading. Water movements in a basin are capable of producing large deflections⁸ which can be coherent at separate localities and a similar effect could arise through changing bedding errors⁹ as a basin infills. Thus observation of between-core coherence arising from these factors may mistakenly lead to a 'within-site reinforcement syndrome'. Thompson and Berglund's three-item criteria must be qualified according to individual circumstances and only a detailed consideration of all factors at each locality can decide the reliability of palaeofield directions. The Stjärnö data survive as important evidence for a Late Weichselian excursion in Sweden which will achieve credibility as a world-wide phenomenon if palaeomagnetic results from sediments reveal global similarities, yet this is the objective which the authors seek to denounce.

MARK NOËL

*Department of Geophysics and
Planetary Physics,
School of Physics,
University of Newcastle upon Tyne,
Newcastle upon Tyne, UK*

¹ Thompson, R. & Berglund, B. *Nature* **263**, 490–491 (1976).

² Noël, M. *Geol. Förel. Stockh. Förh.* **97**, 357–367 (1975).

³ Aitkin, M. J. *Phil. Trans. R. Soc.* **269A**, 77–88 (1970).

⁴ Keen, M. J. *Deep-Sea Res.* **10**, 607–622 (1963).

⁵ Noël, M. thesis, Univ. Newcastle-upon-Tyne (1976).

⁶ Noël, M. *Trans. Am. geophys. Un. (EOS)* **57**, 657 (1976).

⁷ Berglund, B. *Geol. Förel. Stockh. Förh.* **93**, 11–45 (1971).

THOMPSON AND BERGLUND REPLY—Noël¹ assumes that precise palaeomagnetic data are an accurate record of the ancient geomagnetic field. We distinguish between precision and accuracy, and stress the importance of within-site repeatability as one requirement for the recognition of accurate palaeomagnetic recording of geomagnetic excursions.

Noël questions the reliability of our Björkeröds Mosse data² because of his qualitative assessment of the high variability of the sediments and scatter of palaeomagnetic directions. Inspection of the palaeochemical diagrams (Fig. 1 of ref. 3), rather than simply the column of layer numbers, shows that between ~13,000 and 11,000 BP the rate of sedimentation varied from only 50 to 100 mm per 100 yr, the sediment was never coarser than clay grade and that certain of the 'sedimentary units' alluded to by Noël are in fact transition zones. The mean direction of the oriented, cleaned, normally magnetised subsamples spanning the period ~13,000–11,000 BP is dec.=2°, inc.=66°, giving unit weight to each core, with an average 95% cone of confidence of radius of only 3.5° for each core section. We cannot agree with Noël's summary that these data are from sedimentologically complex cores with high noise content.

Noël stresses that "The Revised Swedish Time Scale also provides a precise chronology" for varved clays. Most Swedish Quaternary geologists are now convinced that there is no revised absolute time scale—based on varves—available for South Sweden. There seems to be a discrepancy of 600–800 yr between the varve years and the radiocarbon years south of the Central Swedish Moraines (B.B., in preparation). This is partly due to complications in the varve chronology, partly due to ¹⁴C variations in the atmosphere. Noël consequently uses a floating chronology which is less accurate than the biostratigraphic zones and transferred chronozones of the Björkeröds Mosse sections.

Noël misrepresents our conclusions. Correctly stated they are, "The Swedish late Weichselian reversal has become an example of Watkins's reinforcement syndrome" and "Furthermore, reversals or excursions of the geomagnetic field from other parts of the

world around this time should not be dated or correlated with the earlier Swedish results".

Finally, Noël incorrectly discusses our three minimum requirements as though they were absolute criteria. We reiterate that our criteria are to be regarded as minimum practical criteria for recognising geomagnetic excursions in recent sediments and that they represent only the first of several hurdles which a geomagnetic excursion must pass before gaining recognition as a global phenomenon.

Fairbridge⁴ has recently suggested the Late Weichselian geomagnetic 'reversal' is correlated with a short sharp cold fluctuation in climate. He discusses a mechanism by which climate could be controlled by geomagnetic changes. We would explain such a correlation by an opposite mechanism, namely that sudden climatic changes produce variable sediments which increase the probability of spurious palaeomagnetic data and hence apparent geomagnetic changes.

R. THOMPSON

Department of Geophysics,
Edinburgh University,
Edinburgh, UK

B. BERGLUND

Laboratory of Quaternary Biology,
Institute of Geology,
Lund University, Sweden

¹ Noël, M. *Nature* 267, 180 (1977).

² Thompson, R. & Berglund, B. *Nature* 263, 490-491 (1976).

³ Berglund, B. E. & Malmer, N. *Geol. Förr. Stockh. Förh.* 93, 575-589 (1971).

⁴ Fairbridge, R. W. *Nature* 265, 430 (1977).

Liquid nitrogen enhancement of partially annealed fission tracks in glass

PILIONE and Gold¹ described a means of revealing partially annealed fission tracks in glass by immersing the glass (NBS SRM 962)² into liquid nitrogen. According to their discussion it was possible to increase the total number of etchable tracks by increasing the immersion time of the glass in liquid nitrogen. Using this information we attempted to duplicate this work using the same NBS SRM 962 glass.

In a preliminary experiment, we immersed a piece of SRM 962 glass wafer into liquid nitrogen for approximately 8 h. Before the immersion the

glass had been heated to 230 °C for 1 h such that 57% of the tracks were annealed and then the glass was polished. This piece of glass and an unannealed piece, which had previously been broken from the same SRM 962 wafer, was etched in 16% HF at 23 °C for 75 s. Then the fission tracks in both pieces were counted and the track densities compared. The track counting statistics, however, were such that a definitive conclusion could not be drawn.

The experiment was repeated using several pieces of the SRM 962 glass and different continuous liquid nitrogen immersion times. At the end of each given immersion time, pieces of the glass were removed, allowed to reach room temperature, polished, and etched in HF as before. The results of this experiment are summarised in Table 1. We found no significant change with respect to the increased number of etchable tracks in the glass after its immersion in liquid nitrogen.

From our results we conclude that liquid nitrogen has no effect on the annealed tracks in glass. Subsequent discussions with Pilione and Gold have indicated that this effect was only observed when the glass had been polished before exposure to nitrogen. Therefore our conclusion is that the enhancement of partially annealed tracks is perhaps a surface phenomenon and has no effect on the interior matrix of the glass.

G. A. WAGNER

Max-Planck-Institut für Kernphysik,
69 Heidelberg, FRG

B. S. CARPENTER

National Bureau of Standards
Washington, DC 20234

¹ Pilione, L. J. & Gold, D. P. *Nature* 262, 773 (1976).

² Carpenter, B. S. & Reimer, G. M. *NBS Spec. Publ.* 260, 49 (1974).

PILIONE AND GOLD REPLY—It is apparent from discussions between Pilione and Carpenter that different criteria are used for the identification of fission tracks in glass. But any 'track identification' difference should not affect the results if the density of 'artefacts' remained constant with immersion in liquid nitrogen, because the values reported were normalised in ratio form. We are investigating the effects of

chemical etchants on fission tracks and 'artefacts' in glass, in order to resolve this dilemma. We intend to report in the near future on recent experiments indicating the enhancement effect we noted earlier¹ is probably a surface rather than a volume phenomenon².

L. J. PILIONE

Department of Physics, Altoona
Campus,
Pennsylvania State University,
Altoona, Pennsylvania 16603

D. P. GOLD

Department of Geosciences,
Pennsylvania State University,
University Park, Pennsylvania 16802

¹ Pilione, L. J. & Gold, D. P. *Nature* 262, 773 (1976).

² Wagner, G. A. & Carpenter, B. S. *Nature* 267, 182 (1977).

Lack of correlation between tight junction morphology and permeability properties in developing choroid plexus

MOLLGARD *et al.*¹ report on morphological study of the sheep choroid plexus during foetal development. Excellent freeze-fracture replicas clearly show no difference in the ultrastructural features of the tight junctions in the choroid epithelium between 40- and 125-d foetuses. The authors reasonably agree that changes in tight junction strand number cannot account for alterations in permeability properties during development, since changes in strand number do not occur.

The reader unfamiliar with the blood-brain-cerebrospinal fluid system is likely to be left with the impression that these results disprove the existence of a correlation between tight junction morphology and permeability in the choroid epithelium itself and therefore in epithelia in general. This is not so. All the permeability studies²⁻⁴ referred to were made in intact animals or foetuses, in which transport of a solute between blood and cerebrospinal fluid can occur indirectly through the walls of the cerebral vessels and the interstitial fluid of brain as well as directly across the choroid epithelium. The observed developmental alterations in permeability with respect to cerebrospinal fluid are probably partly, and perhaps wholly, due to maturation of the cerebral vessels, that is, of the blood-brain barrier^{2,3}. Unless the authors or others can as clearly show for the cerebral endothelial cells of the sheep what has been so well demonstrated for the choroid epithelium, namely that there is no morphological change in the tight junctions of the blood-brain barrier from 40 d gesta-

Table 1 Effect of liquid nitrogen immersion on partially annealed fission tracks in glass

Glass condition	Track density* (tracks mm ⁻²)	Track retention
Not annealed	402.3	100%
Annealed, 230 °C for 1 h	231.4	57.5%
No N ₂ treatment	231.0	57.4%
1 d in N ₂	228.2	56.7%
7 d in N ₂	223.7	55.6%
30 d in N ₂		

*Counting statistics (1σ) is ±3%.

tional age to full development, no unequivocal conclusions can be made concerning structure and function in the tight junctions of the choroid plexus from the recorded observations.

M. W. BRADBURY

Department of Physiology,
King's College,
Strand,
London WC2, UK

¹ Møllgård, K., Malinowska, D. H. & Saunders, N. R.

Nature 264, 293–294 (1976).

² Bradbury, M. W. B. et al. *J. Physiol., Lond.* 227, 591–619 (1972).

³ Evans, C. A. N., Reynolds, J. M., Reynolds, M. L., Saunders, N. R. & Segal, M. B. *J. Physiol., Lond.* 238, 371–386 (1974).

⁴ Dziegielewska, K. M. et al. *J. Physiol., Lond.* 258, 86P–87P (1976).

MØLLGÅRD *et al.* REPLY—We thank Dr Bradbury for raising an important aspect of the interpretation of our work which was insufficiently dealt with in our letter because of lack of space. In parallel with our studies of the choroid plexus and cerebrospinal fluid (CSF) we have also studied the ultrastructure of cerebral vessels and penetration of the same markers (non-electrolyte and plasma proteins) into developing brain. We have previously published results of ultrastructural studies of vessels in the immature brain¹. These results, although preliminary, suggested that interendothelial cell tight junctions develop very early in gestation. This has recently been confirmed (unpublished observations) in sheep brain as early as 30 d gestation. The strand number seems to be greater than in the choroid plexus at all ages so far examined. It is technically more difficult to obtain freeze-fracture replicas of cerebral endothelial tight junctions, so it is doubtful if it will be possible to provide evidence for endothelia on the scale already obtained for the epithelial cells of the choroid plexus. However, all endothelial cell junctions so far examined, either by freeze-fracture or by thin section, have had the appearance of well formed tight junctions.

Our permeability data (ref. 2 and unpublished) in 60-d sheep foetuses show that there is more restriction on penetration of protein from blood into brain than into CSF. This implies the presence of junctions which are functionally tight to protein. The data on foetal sheep electrolytes in CSF and plasma³, mentioned by Dr Bradbury, showed that CSF–plasma gradients were present for chloride and magnesium by 60 d gestation. This probably indicates the presence of both the appropriate ion pumps and a permeability barrier adequate to prevent loss of the gradients. Our non-electrolyte

data (ref. 44 and unpublished) suggest that inulin and sucrose penetrate into brain and CSF to a similar degree if the brain extracellular space is assumed to be equivalent to the chloride space. It may be, however, that these materials reach the brain in early development predominantly through the choroid plexus since this is disproportionately large in the immature foetus and the brain is poorly vascularised.

K. MØLLGÅRD

Institute of Medical Anatomy A,
University of Copenhagen,
Universitetsparken 1,
2100 Copenhagen, Denmark

D. H. MALINOWSKA

N. R. SAUNDERS

Department of Physiology,
University College London,
Gower Street, London WC1, UK

¹ Møllgård, K. & Saunders, N. R. *J. Neurocytol.* 4, 453–468 (1975).

² Dziegielewska, K. M. et al. *J. Physiol., Lond.* 285, 86–87P (1976).

³ Bradbury, M. W. B. et al. *J. Physiol., Lond.* 227, 591–610 (1972).

⁴ Evans, C. A. N., Reynolds, J. M., Reynolds, M. L., Saunders, N. R. & Segal, M. B. *J. Physiol., Lond.* 238, 371–386 (1974).

Collagen microfibrils and cell attachment proteins

PEARLSTEIN¹, following Klebe² in the search for a cell attachment protein (CAP) (ref. 3) between mesenchymal cells and collagen, seems to disregard the lesson learnt from Jamieson's similar work on platelets⁴. They omitted to examine the microfibrillar structure of their reconstituted collagen substrata^{1–4}. Pearlstein¹ treats his collagen with 8 M urea to minimise 'non-specific' adhesion (see below). But he fails to note that urea inhibits fibrillogenesis⁵, or that collagen-binding both by platelets^{5,6} and fibroblasts⁶ depends crucially on microfibril aggregation, which also has a key role in the adhesion of BHK cells to a microfibril-forming polysaccharide substratum^{7,8}.

Moreover, the divalent ion requirements of CAP^{1,2} seem to conflict with those of other experiments^{9,10} which do not require CAP for adhesion to various forms of collagen. Rabinovitch and DeStefano showed that Mg or Mn by themselves promote adhesion of carcinoma cells to gelatin⁹, and I have confirmed this for BHK cells on adsorbed films of soluble collagen¹⁰.

Therefore, the fact that aggregation and divalent cations, respectively, can promote adhesion independently of

CAP calls for re-assessment of the "background" of "non-specific" adhesion (that is, in the absence of CAP) subtracted by Pearlstein¹ and Klebe^{2,3}. I suggest that their background adhesion, far from being an inconsiderable artefact, may actually represent a physiologically significant attachment of the mesenchymal cells to aggregated, microfibrillar parts of the reconstituted collagen substrata, under the influence of divalent cation alone; while their CAP-specific adhesion may be to a single-stranded (for example, urea-treated) form of collagen, not found in intact connective tissue. So far, unfortunately, CAP does not seem to have been tested on a physiologically intact form of collagen, such as omentum.

Hence the experiments on mesenchymal cells^{1–3} should be repeated as those on platelets⁴ have been, that is, with confirmed microfibrils^{5,6}. Moreover, since there is evidence that various mesenchymal cells^{7,8,10–12} explore micro-architectural and mechanical aspects of their substrata, critical work with mesenchymal cells on collagen should also describe the size, spatial distribution and rigidity of microfibrillar particles.

N. G. MAROUDAS

71 Park Avenue North,
London NW10, UK

¹ Pearlstein, E. *Nature* 262, 497–500 (1976).

² Klebe, R. J. *Nature* 250, 248–251 (1974).

³ Kleinman, H. J., McGowan, E. B. & Klebe, R. J. *Biochem. biophys. Res. Commun.* 72, 426 (1976).

⁴ Jamieson, G. A., Urban, C. & Barber, A. J. *Nature* 234, 5 (1971).

⁵ Barnes, M. J., Gordon, J. L. & MacIntyre, D. E. *Biochem. J.* 160, 647–651 (1976).

⁶ Gordon, J. L. & Dingle, J. T. *J. Cell Sci.* 16, 157 (1974).

⁷ Maroudas, N. G. *Nature* 254, 695–696 (1975).

⁸ Maroudas, N. G. in *Supramolecular Structure: Cell Shape and Surface Architecture* (eds Henning, U. & Revel, J. P.) (Liss, New York, 1977).

⁹ Rabinovitch, M. & DeStefano, M. J. *J. Cell Biol.* 59, 165 (1973).

¹⁰ Maroudas, N. G., *J. cell. Physiol.* (in the press).

¹¹ Albrecht-Buehler, G. *J. Cell. Biol.* 69, 275–286 (1976).

¹² Dunn, G. A., and Heath, J. F., *Expt Cell Res.*, 101, 1–14 (1976).

Stereospecificity and clinical potency of neuroleptics

ENNA *et al.*¹ have drawn attention to the potential, and some of the difficulties, of studies of the stereospecificity of neuroleptic drugs in elucidating the anti-schizophrenic action of these compounds. Although they have provided new data of considerable value we feel that they have obscured one issue. There is an increasing body of evidence on the stereospecificity of these drugs in *in vitro* assay systems and some evidence on their effects on animal behaviours, but as far as we are aware there are no systematic studies

of the stereospecificity of the anti-psychotic effect in man. Illnesses comparable with schizophrenia have not been identified in animals although some drug-induced models are available.

Enna *et al.* state that " α -flupenthixol and *cis*-thiothixene are much more potent clinically and in animals than their geometrical isomers β -flupenthixol and *trans*-thiothixene respectively. Similarly (+)-butaclamol is much more potent than its optical isomer (–)-butaclamol". These statements are supported by references which include data from animal behavioural experiments but no clinical findings. Therefore there seems to be a risk that the argument will be accepted that data obtained from an animal model of psychosis (for example amphetamine-induced abnormal behaviours) can provide definitive information concerning the nature of schizophrenia or its response to drugs. We suggest that such information can only be obtained from studies on patients with the disease, although this is difficult and time consuming.

If such studies can be designed to include not only a comparison of one isomer with another but of both with a placebo-treated group they can, however, in spite of the difficulties emphasised by Enna *et al.*, yield information beyond that relating to the *in vitro* receptor systems recently studied. If, for example, it can be shown that α -flupenthixol is clinically more active than the β -isomer and the latter is more active than placebo we may conclude on the basis of the data of Enna *et al.* not only that blockade of the dopamine, and possibly 5-HT, receptor may be relevant to the mechanism of action but that another action is also involved. If, on the other hand, the α -isomer is potent and the β -isomer is no more active than placebo, it seems safe to conclude that actions on those receptors on which the two isomers have been shown to be equally active (for example cholinergic, adrenergic and opiate receptors) are not relevant to the therapeutic effect.

T. J. CROW
J. W. F. DEAKIN
E. C. JOHNSTONE

MRC Clinical Research Centre,
Watford Road, Harrow,
Middlesex, UK

¹ Enna, S. J., Bennett, J. P., Burt, D. R., Creese, I. & Snyder, S. H. *Nature* 263, 338–341 (1976).

ENNA *et al.* REPLY—Crow *et al.*¹ rightly pointed out the importance of clinical studies in establishing for certain whether drugs, such as isomers of

Table 1 Neuroleptic stereoselectivity at α -noradrenergic receptor binding sites in rat brain

	Affinity for α -receptors	
	K_i (nM)	Relative stereoselectivity
(+)-Butaclamol	35	70-fold
(–)-Butaclamol	2,500	
<i>cis</i> -Thiothixene	6.6	
<i>trans</i> -Thiothixene	150	23-fold
α -Flupenthixol	10	
β -Flupenthixol	36	3.6-fold

α -Receptors in membranes from whole rat brain were labelled by binding of the potent α -antagonist ³H-WB-4101 (2-([2', 6'-dimethoxy] phenoxyethylamino) methyl benzodioxan) subtracting as blank values binding in the presence of 100 μ M (–)-noradrenaline. Inhibition constants (K_i) were determined by measuring the effects of 3–4 concentrations of each drug in triplicate.

neuroleptics, possess anti-schizophrenic activity. It is commonly assumed that neuroleptics act therapeutically in schizophrenic patients by blocking dopamine receptors so that the isomer which fails to act on dopamine receptors would be anticipated to lack clinical effectiveness. This is only a presumption, which must be directly examined in clinical studies. If supposedly inactive isomers have therapeutic effect, one might assume that they exert part of their action through a system which does not display stereospecificity.

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

The focus of our paper² was that some neuroleptics are stereospecific at 5-HT as well as dopamine receptors so that a stereospecific pharmacological action does not necessarily involve only dopamine synapses. Recently we have succeeded in identifying α -noradrenergic receptor binding in brain membranes using an agonist ³H-clonidine or antagonist ³H-WB-4101 (refs 3,4). Several neuroleptics have very high affinities for the α -noradrenergic receptor, often similar to their affinities for dopamine receptors⁵. But, the relative potencies of neuroleptics differ in competing for α receptors and dopamine receptors. Affinity for α receptors parallels closely pharmacological tests of α -noradrenergic blockade and tends to correlate with the ability of neuroleptics to elicit

sedation and orthostatic hypotension⁵. We have now found that neuroleptics display marked stereoselectivity at α -adrenergic receptor sites. The extent of stereoselectivity for the butaclamol isomers is as great at α -noradrenergic receptors, as at dopamine receptors (Table 1). The potency of *cis*-thiothixene at α -noradrenergic receptors is similar to its potency at dopamine receptors, though the 23-fold stereoselectivity is somewhat less than that at dopamine receptors. The extent of stereoselectivity of flupenthixol isomers at α receptors is substantially less than at dopamine receptors. As some clinical improvement in psychiatric patients, including schizophrenics, could be associated with 'tranquilizing' effects secondary to α receptor blockade, these receptors and their stereoselective interactions with neuroleptics represent yet another neurotransmitter system which should be taken into account in explaining neuroleptic pharmacology.

S. J. ENNA

Departments of Pharmacology and
Neurobiology and Anatomy,
University of Texas Medical School
at Houston,
PO Box 20708,
Houston, Texas 77025

J. P. BENNETT, JR
D. R. BURT
I. CREESE
D. C. U'PRICHARD
D. A. GREENBERG
S. H. SNYDER

Departments of Pharmacology and
Psychiatry,
Johns Hopkins University,
School of Medicine,
Baltimore, Maryland 21205

¹ Crow, T. J., Deakin, J. F. W. & Johnstone, E. C. *Nature* 267, 183–184 (1977).

² Enna, S. J., Bennett, J. P., Burt, D. R., Creese, I. & Snyder, S. H. *Nature* 263, 338–341 (1976).

³ Greenberg, D. A., U'Prichard, D. C. & Snyder, S. H. *Life Sci.* 19, 69–76 (1976).

⁴ U'Prichard, D. C., Greenberg, D. A. & Snyder, S. H. *Molec. Pharmac.* (in the press).

⁵ Peroutka, S. J., U'Prichard, D. C., Greenberg, D. A. & Snyder, S. H. *Neuropharmacology* (in the press).

reviews

History of African crops

A. H. Bunting

Origins of African Plant Domestication. Edited by J. R. Harlan, J. M. de Wet and A. B. L. Stemler. Pp. xiii+498. (Mouton: The Hague and Paris; Aldine: Chicago, 1976.) Dfl.84.50; \$32.50.

THIS book contains the papers of the 58th Wenner-Gren Symposium on the origin of African plant domesticates, held at Burg Wartenstein in August, 1972. It opens with a series of three general papers, on the origins of domestication and of agriculture in Africa (Harlan, de Wet, Stemler), on causes (Pfeiffer) and on the problems of philosophy, technique and evidence (Higgs). Between them, they spell out the central question—at what periods, where, and why did African people begin deliberately to plant, tend and harvest some of the indigenous food and other useful plants of Africa; and how far, in doing these things, were they affected by influences from Asia? Over these topics controversy has smouldered for many years, alongside the more general endeavour to describe, more satisfactorily, the nature of the Neolithic transition and the emergence and spread of agriculture.

The climatic and ecological scene is set by van Zinderen Bakker. During the last glacial period, up to about 12,000 yr ago, Africa was dry, so that the Sahara spread as far south as 13°N (not S, as an oversight suggests). In a subsequent more or less warm and wet period, cattle, crocodiles and men could inhabit much of what is now desert, and the internal delta of the White Nile reached almost as far north as the modern junction of the Niles at Khartoum. This gave way about 4,500 yr ago to a cooler and drier period which may have driven Saharan man from an Eden of hunting, gathering and fishing into the drudgery of agriculture.

Seven archaeological and cultural/historical papers follow. Three are general (Clark, Shaw, P. E. L. Smith) and four more specialised (Munson, Flight, David, Wendorf and Schild). The evidence comes from a regrettably small number of sites, and there is not yet enough of it to tell a continuous or coherent story. Wheat and barley

from Asia were grown in Egypt more than 7,000 yr ago. It is easy to be convinced that indigenous *Pennisetum*, after being gathered for several thousand years, had given mankind a new crop, bulrush millet, by about 2,500 yr ago. Sorghum and cowpeas appear in West, East and Southern Africa in archaeological contexts, none of which is more than 1,500 yr old. From the Ethiopian plateau there is only one early record, of barley and chickpea, at about 2,500 yr ago.

The last seven papers are on botanical and agricultural topics. Purselove presents a useful compendium on the crops domesticated in Africa, the history of exchanges between Africa and Asia, and the entry of American species. An essay by Harris on the origins of agriculture in West Africa is based too firmly for me on geographical and ecological 'first principles', but many will find in it, as its author hopes, useful hypotheses to be tested against observations, new and old. Scudder's brave reconstruction of prehistoric land-use in Zambia is illuminated by studies of modern agriculture in the same region. Then the botanists have the last words—on yams (Coursey) and African cereals: sorghum (two papers by de Wet, Harlan, Price and Stemler) and other species (by the late Professor R. Portères, a valuable English summary of over thirty years of devoted and scholarly labour). The two sorghum papers base a geographical history of the cultivated sorghums, starting with initial domestication in the eastern part of the sub-Saharan region, on a range of modern biosystematic and agricultural criteria, but in the absence of early archaeological records the timescale has to depend on the records of sorghum in India 3,000 yr ago.

This book is valuable because it assembles much of what is known, what is said to be known, and what is thought at the present time about the history of African crops. It is strongly suggested that the chief African cereals, and the domesticated yams, emerged without benefit of influences from outside Africa. But the timescales are tantalising: most if not all of the known history happened after agri-

culture had been established in the Nile Valley, and long after it was the basis of a way of life only 400 km away at Jericho and of course further north and east. It is difficult to imagine that the news did not get around. Man, including African man, is far too mobile for that. One wonders whether African oral tradition, difficult though it may be to date, has not a contribution to make here.

It must be important to learn more about human movement in antiquity: the apparently ancient navigation of the Indian Ocean from China to Ceylon, the Gulf, Suez and East Africa, the links across the Red Sea between the similar agricultures of the highlands of Yemen and of Ethiopia, the east-west traffic through what are now the Sahelian and savanna zones from the Blue Nile region (where every *jebel* seems to have been a workshop in days gone by) to the Atlantic; the north-south traffic across what is now the desert from the Mediterranean and the Nile, along the routes—like the *Darb al Arba'in*, the road of forty days—to Kordofan, Darfur, the Tibesti and West Africa (which some suggest brought the builders of stone circles to Senegal and the Gambia, although we know not when).

Then I am tantalised by querns and mortars, the two distinct means by which African people grind grain. The quern, an ancient device in Asia, seems to have spread in Africa before agriculture; the mortar seems to be African. Are the crops and the means for preparing them historically related, and if so, how?

Finally a question about vegetation. Chevalier divided the different sorts of vegetation he saw in West Africa into the forest, savannah and other zones. These correspond broadly with current climates, and many people seem to regard them as 'natural' biogeographical divisions. Yet without exception they are artefacts of agricultural man. What did West Africa look like in the wetter and warmer period 6,000 yr ago, when farming was at most in its infancy? □

A. H. Bunting is Professor of Agricultural Development Overseas at the University of Reading, UK.

Distinguished Arctic explorer

Sir John Richardson: Arctic Explorer, Natural Historian, Naval Surgeon. By R. E. Johnson. Pp. xii+209+33 plates. (Taylor and Francis: London, 1976.) £15.

It is remarkable that this should be the first attempt at a full-length biography of Sir John Richardson since the account by his nephew, John McIlraith, in 1868.

It comes as no surprise to learn that the man who distinguished himself in Arctic exploration, and natural history and introduced reforms in medical practice, was a precocious youth. Apprenticed to a surgeon at the age of thirteen, Richardson began his medical studies at the University of Edinburgh the following year. A little before his twentieth birthday he joined the Royal Navy as an assistant surgeon.

The direction of his life was decided when he was made surgeon and naturalist on an overland expedition to the Canadian Arctic Ocean under Lieutenant John Franklin in 1818. During the next three years this expedition covered 5,500 miles in appalling conditions, losing eleven men, one of whom Richardson was forced to shoot in self-defence. Richardson joined Franklin again as surgeon and naturalist on his second Arctic expedition in 1825, and when Franklin failed to return from his third expedition in 1845, Richardson at the age of sixty-one set out for the Arctic once more in a vain attempt to find his friend.

Sorry, for copyright reasons some images on this page may not be available online

Richardson's scientific career began with his zoological and botanical appendices to Franklin's and Parry's accounts of their Arctic expeditions. Sir W. J. Hooker's *Flora Boreali-Americana* (1840) was based principally on the plant collections of Richardson and his very able assistant, Thomas Drummond. His scientific reputation became firmly established through his contributions to the four-volume *Fauna Boreali-Americana* (1829-1837), a pioneer work in Arctic biology.

During his time as physician at the Royal Naval Hospital, Gosport, he created a museum of comparative anatomy and taxonomy which was consulted by scientists as eminent as

Darwin, Lyell, Hooker and Gray. Richardson was a prodigious worker, as the complete bibliography of his books and periodical articles in this book testifies.

Very little seems to have escaped the author's diligent searching for every scrap of information about Richardson. Even the translated text of Richardson's thesis for his doctorate in medicine, *De Febre Flava* (1816), has been included. Research workers in this field will find his list of the main collections of Richardson's letters useful. One of the appendices lists all the known portraits of Richardson followed by the species of animals named by or after him, but through some oversight the plants have been omitted.

In a work that is so full of facts some errors are inevitable: Richardson was elected a Fellow of the Royal Society in 1825, not 1835 (p65); J. E. Gray retired as Keeper of the Department of Zoology at the British Museum in 1874, not 1872 (p114); Sir W. J. Hooker was appointed Director of Kew Gardens in 1841, not 1847 (p114).

Dr R. E. Johnson has methodically assembled all the data necessary for a definitive life of Sir John Richardson. But Richardson does not come alive through his pen, and there is no real assessment of a naturalist who was also one of the outstanding ichthyologists of the nineteenth century.

R. Desmond

R. Desmond is Deputy Keeper, India Office Library and Records, and President of the Society for the Bibliography of Natural History.

Process of language understanding

Computational Semantics: An Introduction to Artificial Intelligence and Natural Language Comprehension. Edited by E. Charniak and Y. Wilks. Pp. vii+294. (North-Holland: Amsterdam, New York and Oxford, 1976.)

"COMPUTATIONAL SEMANTICS", say Charniak and Wilks in the preface to their book, "is not so much a new subject as a new way of looking at old questions—those concerning meaning, language and understanding. It is based on the assumption that a good way to explicate such difficult notions is to work toward the programming of an automaton, or digital computer, so that it could be said to understand language." This assumption characterises much recent research in Artificial Intelligence (AI), and it is the AI viewpoint which is strongly put forward by this book.

To demonstrate the significance of this viewpoint, Charniak, Wilks and the other contributors to this volume have placed computational semantics within the context of related language disciplines, providing a broad overview of linguistics, the psychology of language and memory, the philosophy of language, AI approaches to language comprehension, and fundamentals of computer programming. In bringing together all of these cognitive science topics in a tutorial format, the editors have taken an important step towards interdisciplinary consciousness raising. Such an ambitious step is not without flaws, of course—the most serious being the use of heavy attacks on straw men within generative semantics and psychology in order to glorify AI approaches to language comprehension. The drawing of these 'battle lines' detracts from an otherwise noble cause.

The book stands strongly on its positive treatment of tother topics—for instance, it provides an excellent survey of AI language comprehension programs, particularly in the second of Charniak's

two chapters on inferencing and the second of Wilks' two chapters on parsing. Both of these chapters provide clear accounts of recent AI research on language understanding, and suggest some new dimensions along which various AI programs may be compared.

Students interested in theories of language, meaning, and understanding would be well advised to read this book, and to develop a familiarity with all of the areas it covers. Psychologists, linguists and philosophers will no doubt find fault with the sections dealing with their own specialities. If they keep an eye out for straw men, however, they will find that the other sections of the book provide a fine background to some of the most exciting contemporary research into the processes of language understanding.

Marc Eisenstadt

Marc Eisenstadt, formerly a Research Fellow at the Department of Artificial Intelligence, University of Edinburgh, is now Lecturer in Psychology at The Open University, Milton Keynes, UK.

Major histocompatibility complex

Histocompatibility. By George D. Snell, Jean Dausset and Stanley Nathenson. Pp. xiv+401. (Academic: New York, San Francisco and London, 1976.) \$29.50; £18.

THIS is an excellent book for seasoned professionals and advanced students with a substantial background in mammalian genetics and immunology. Emphasis is on the much studied, major histocompatibility complex (MHC) of mice (*H-2*) and men (*HLA*). It is difficult to find fault with this masterfully written treatise. The coverage is thorough and well-documented. The subject index is admirably detailed. The list of abbreviations given at the outset is helpful in coping with inevitable acronyms. Because so much of the current literature deals with the findings of the past few years, it is also refreshing to encounter each theme in historical perspective. One only wishes the scope extended to non-mammalian vertebrates and invertebrates, but this is newer territory awaiting exploration in depth. This superbly comprehensive volume will be the definitive reference for histocompatibility characteristics of mice and men for years to come. If progress continues at the present rate, however, an updated edition will surely be desired about 1980. The principles so nicely explained in this book should of course remain valid, but the details will surely become more complex.

The first eight chapters take the reader progressively from essential principles of transplantation immunogenetics through allogeneic polymorphism, *H-2* complex loci, and immune response (*Ir*) genes in mice. In chapters 9, 10 and 12, the serology, genetics, and disease associations of the *HLA* complex in man are described in a cogent manner. Chapter 11 deals comparatively with the biochemistry of *H-2* and *HLA* alloantigens. The predominant role played by the MHC in cell interactions, especially lymphocyte interactions such as mixed lymphocyte reactions, cell-mediated lysis, T-B lymphocyte collaboration, and hybrid resistance in mice is admirably presented, especially in chapter 8.

At any given time, every field has a prevailing dogma and histocompatibility or transplanatation immunogenetics is no exception. Although emphasis is on the MHC as the only gene complex imposing a "very strong barrier to transplants", the roles of other *H-* or *Ir* genes are trenchantly considered in chapters 3, 7, and 13, the final chapter. Even recipients of *HLA*-identical organ transplants will usually reject such allografts promptly in the absence of drug immunosup-

pression; immunodeficient recipients of allogeneic bone marrow regularly succumb to fatal graft-versus-host disease despite identity for all four *HLA* loci. Other *H* systems then may often have 'major' impact. Moreover, as the authors aver in several chapters, slight pre-immunisation with donor-type alloantigens may convert a weak non-MHC barrier into a strong one. Uniqueness of the MHC, if indeed each species has only one such *main system* (book italics), apparently resides in the contiguity of the constituent loci and their interaction. Many of the constituent *H-2* loci taken singly in tests between recombinant congenic lines have effects similar to moderate non-*H-2* loci. In light of other recent evidence, much of the mammalian genome may be organised into large functional units with relevant developmental and regulatory information associated with component structural genes.

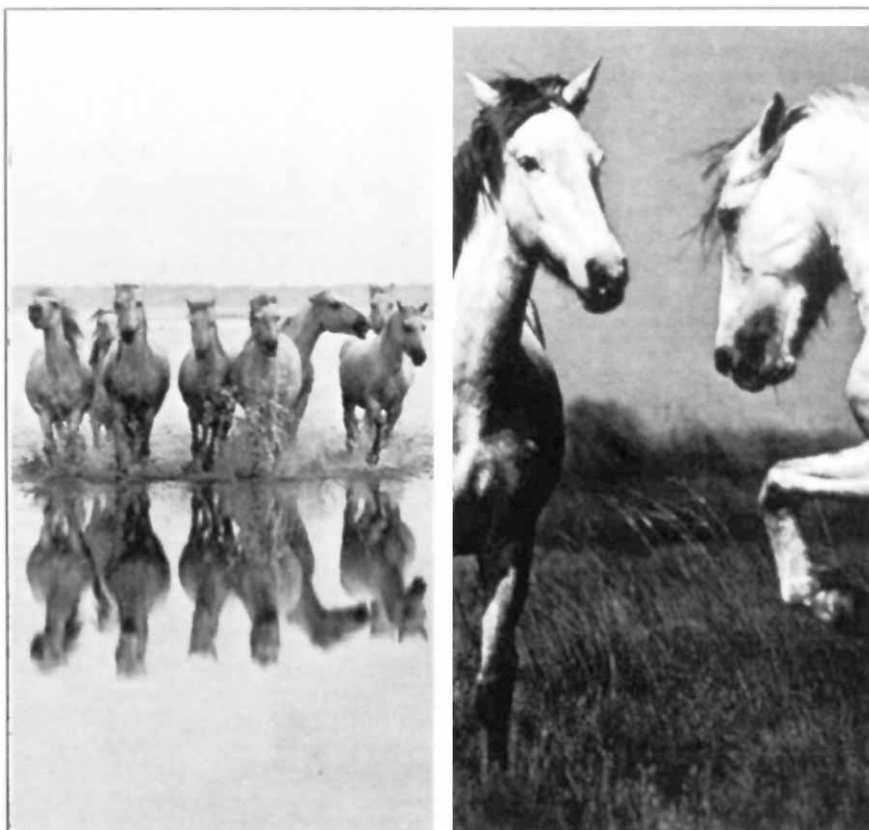
Snell, Dausset and Nathenson deal perceptively with discordant findings. For example, divergent data concerning the genetic 'laws' of transplantation, sources of specific immune blocking, and the chemistry of MHC antigens are evaluated in a commendably evenhanded manner. Moreover, they suggest further experiments to resolve ambiguities. An understanding of sequential *HLA* gene-antigen relationships in man has now become difficult for non-specialists. An illustrative diagram proceeding from the chromo-

some No. 6 loci→ alleles and antigens→ haplotypes→ phenotype would have been useful in chapter 9 to epitomise the immunogenetics of the *HLA* system.

Chapter 7 deals concisely with immune response genes and probable mechanisms of *Ir* gene action in mice. This theme is pursued again in man in an incisive chapter 12 on histocompatibility genes and disease—especially associations between *HLA* and disease susceptibilities. The authors favour the hypothesis that *Ir* and metabolic genes closely linked to the *HLA* genes are actually responsible for susceptibility or resistance. Although acknowledging the undoubted relevance of the *HLA* complex to medical genetics, the reader should retain reservations toward the 'MHC controls everything' viewpoint by remembering that many known *Ir* genes are not linked to the MHC; that only a few human diseases (all non-neoplastic) reveal a strong correlation with particular *HLA* specificities; and that animals below the phylogenetic level of advanced amphibians or bony fishes cope quite well immunologically without having evolved a mammalian-type MHC. Altogether, this volume is an outstanding contribution to the field of transplantation genetics and immunobiology.

W. H. Hildemann

W. H. Hildemann is Professor of Immunogenetics and Dental Research Institute Director at the University of California, Los Angeles, California.



Illustrations taken from *Horses of the Camargue*. Photographs and text by Hans Silvester. Preface by Konrad Lorenz. Pp. 23 + 70 pages coloured photographs. (Phaidon: Oxford, 1976.) £12.50.

Marine pollution handbook

Marine Pollution. Edited by R. Johnston. Pp. 729. (Academic: London and New York, 1977.) £23.40; \$51.

MARINE pollution texts are being published with almost monotonous regularity, and the subject is at best duplicated, but more often fragmentary and of poor perspective. The field is a complex one and calls for a multi-disciplinary approach, and it is not without significance therefore that Dr Johnston has chosen to produce a book by persuading colleagues to write chapters on various specialised aspects.

There are eight contributors to a volume of over 700 pages, with material organised into four principal sections. The first, by the editor, provides a broad brush background to the marine pollution scene, including aspects of fisheries biology and a brief review of the major categories of pollutants, their principal pathways in the marine environment and their potential effects. The second concerns a number of topics of importance to an understanding of the pollution problem, including material on dispersion in estuaries and coastal waters, on heavy metal contamination in the sea, on sewage pollution, on oil pollution and on the impact of pollution on seabirds. The third section is devoted to a review of the problems of measuring biological response; a very intractable problem in terms of the natural environment and one only showing signs of being amenable to treatment in terms of human health. The fourth and final section deals with the legal aspects of marine pollution regulation.

Marine pollution is a very emotive subject, and quite opposing views as to its significance are genuinely held by scientists of all disciplines engaged in oceanographic science. It would be surprising, therefore, if any reviewer could wholly agree with conclusions of one contributor, let alone eight. Nevertheless the book is remarkable in its attempt, particularly in the editor's first chapter, to provide an unbiased view of the subject; and the chapters by Bryan on metals and Topping on sewage are admirable and balanced treatments of their subjects, although Bryan's might have benefited from a slightly more quantitative treatment of the dynamics of metal uptake and loss in organisms.

Cowell's chapter on oil covers much familiar ground but is well presented and readable, although unfortunately marred at the end by an ambiguous statement according significance to mercury pollution. The persistence

of mercury could not conceivably create a global pollution problem since manmade input would have to be maintained at present rates for periods commensurate with the residence time of mercury in the oceans, say 10^4 yr, before a significant increase in the worldwide concentrations of mercury in the oceans could be achieved. It is therefore misleading to state that persistence can necessarily be equated with global contamination, although it may be in surface waters if the pollutant is distributed by way of the atmosphere, as in the case of DDT.

The chapter by Bourne on seabirds and pollution makes interesting reading, and he rightly points out that vulnerability of birds at the end of marine food chains, and the conspicuous way in which dead animals are seen on the strand lines of beaches, often provide the earliest indication of unacceptable situations arising at sea.

Muscular exercise

Biomechanics and Energetics of Muscular Exercise. By Rodolfo Margaria. Pp. x+146. (Clarendon: Oxford; Oxford University: London, 1976.) £7.

It is a pleasure to read a book by an expert expounding his life's work, an *auto-festschrift* as it were. Professor Margaria treats the reader to a feast cooked by a master chef, beginning with an historical note in which he himself appears as making a crucial contribution to the lactic acid-oxygen debt controversy. He leads us, by way of his own and his colleagues' work, through the intricacies of the sources of energy for muscular contraction, quantifying the contributions made by oxidation, glycolysis and phosphagen cleavage (the lactic acid and alactic oxygen debt). There follows a very short and somewhat cryptic chapter on some of the cardiorespiratory changes in exercise. Here is the only serious attempt within the book to use SI units; the style and content make one suspect another and less thoughtful hand. The third and final section is entitled "The biomechanics of human locomotion", in which there is a very clear description of how mechanical work may be measured and the interesting conclusions that may be drawn. The author holds out hope for the non-athletic that they too will be able to run at world (Earth) record speeds when the Olympics are held on the Moon.

The book reflects the character of the man; every now and then the

These, of course, are not always related to pollution.

The final chapter by Moore on legal aspects is a most useful review of the subject. Particularly welcome is the clear expression of the view that the function of marine pollution control is not to prohibit entirely the introduction of wastes to the sea, and that indeed the ability of the sea to absorb and assimilate waste is one of its most valuable resources: rather that marine pollution is part of the conflict between the need to release some contaminating materials to the sea and the desire to maximise the use of the marine environment as a source of food and for recreation and transport.

Alan Preston

Alan Preston is Deputy-Director of Fisheries Research in the Ministry of Agriculture, Fisheries and Food based at Lowestoft Laboratories, and Chairman of the UK Marine Pollution Monitoring Management Group.

complex and erudite arguments are interspersed with friendly pieces of practical advice, inviting the reader to relax his mental grip just for a moment. Nevertheless, the book is not for the faint-hearted, the skimmers or the dippers; it is closely argued and requires dedicated concentration. The mathematics, of which there is a good deal, is helpful and yet simple to follow; no-one need fear the book on that account. The author pleads for the adoption of his findings in everyday life—this unfortunately is the *cri de coeur* of many human physiologists! He spoils the case a little, however, by presuming his own degree of expertise in others; how many, for example, would describe his test for maximum anaerobic power as simple? In a world in which rapid communication is producing a tendency to dull uniformity it is interesting to note that schools of thought may still differ on the most practical matters—in Milan the stepping stool reigns supreme for almost the same stated reasons that the bicycle ergometer holds sway in Sweden.

To whom then would this book be useful? Most definitely for the serious and advanced student, but it cannot be unreservedly recommended as a standard text. Although one may overlook the occasional slipshod editing of text and figures, the use of the old units (calories and kilogram-metres), with never a joule or a watt, panders to middle age, prolongs our difficulties and introduces unnecessary confusion for those nearer the starting line.

Rainer Goldsmith

Rainer Goldsmith is Professor of Physiology in the University of London, UK.

obituary

D. H. Menzel

PROFESSOR Donald Howard Menzel, past director of Harvard Observatory, who died in Boston, Mass. at the age of 75 years on December 14, 1976, was born on April 11, 1901 in Florence, Colorado, and educated at the Universities of Denver and Princeton where he received his doctorate in 1924. After several junior appointments of short duration he joined, in 1932, the staff of the Harvard University which remained his academic home until his retirement in 1971. Within his lifetime he took an active part in the transformation of the American scientific scene from its grass-roots stage to the "big science" which it has become—for better or worse—today.

When Menzel left Princeton in 1924, he was a young theoretical astrophysicist—of whom there were not many in the United States at that time—bent on applications of atomic physics to astrophysical situations, and his first interests centred on the physical processes in gaseous nebulae. In collaboration with several students of his who have since attained a renown of their own (and among whom L. H. Aller and J. G. Baker should particularly be mentioned) throughout the 1930s Menzel made important contributions to the subject which are still well-remembered today.

Another field to which Menzel then turned his professional attention, and to which he remained faithful for a major part of his life, was the Sun. The excitement which Bernard Lyot created in the late 1930s with his observations at Pic-du-Midi of the solar corona outside total eclipses, and his films of the activity of solar prominences, is still a vivid memory among the survivors of that heroic epoch. Menzel was the first to transplant these techniques across the Atlantic, and his efforts led to the foundation in 1939 of the High-Altitude Observatory in the Colorado Rockies near Climax, Colorado, at an altitude of over eleven thousand feet—first as an offshoot of Harvard, and later as an independent institution, which only recently closed its doors because of increasing pollution of its environment (caused by nearby mining activities).

Unlike Lyot, who was a genius at improvisation and preferred to perform experiments with his own hands (to the extent of carrying the crucial parts of his coronagraph to the top of Pic-du-



Midi on his back!), Menzel was a theoretician of managerial type, who worked through other people's hands. His field executive in the Climax operation was mainly Walter Orr Roberts, another student of Menzel's who retired only recently from the directorship of the National Centre for Atmospheric Research in Boulder, of which the High-Altitude Observatory eventually became a part.

Then war came (in which Menzel served his country in the uniform of a reserve naval officer) and with it changes in academic and scientific life which were fully reflected also in Menzel's subsequent career. Like many others, after 1945 Menzel never returned properly to resume the career of a scholar, and his activities veered increasingly towards those of an administrator and promoter of science. These latter abilities, which were at a premium during the years of post-war inflation, eventually led to Menzel's appointment in 1954 as Director of the (then) Harvard College Observatory—a position in which he succeeded Harlow Shapley and which he held for the next 12 years.

His illustrious predecessor possessed, among his many talents, an uncanny gift (perfected during his apprenticeship with G. E. Hale) to extract money for astronomy from rich Americans, with which to augment his observatory's current income. In postwar years this goal proved doubly desirable, but no longer attainable in the old ways; for fewer Carnegies were left in the world (income tax saw to that) and, instead, most available funds were to be found in the hands of the Federal Government in Washington. As a result,

Shapley's successor had to seek additional support for his institution, no longer by waiting on the rich, but by lobbying in the "corridors of power" in Washington for grants or contracts from different branches of the Government through their many agencies.

This was an ominous turn of events for the further development of American science; for while rich Americans were (sometimes) willing to make gifts for a good cause with no strings attached and very little paperwork, the new dispensers of Government largesse knew exactly what they wanted and saw to it that they got it. As a result, the Harvard Observatory during Menzel's regime got often involved in work in new fields, such as geophysics, in which it had neither tradition nor sufficient expertise. At times, the Observatory seemed to differ only in location, but not in spirit, from the many research laboratories which sprang up during the "roaring sixties" along Route 128 around Boston as offshoots of Harvard or M.I.T. to serve the customer rather than science—at a time when men were heading for the Moon, and when nothing seemed impossible on Earth.

It is remarkable that, in this atmosphere and with so many administrative preoccupations, Menzel found also time to write or edit a considerable number of books (by himself, or in collaboration with others), ranging from collections of mathematical formulae to polemics about UFOs. Some, like his popular book on the Sun, enjoyed a well-earned popularity in their time, but none seems to have left any deeper mark on contemporary science.

Needless to say, a life so busy has its drawbacks for institutions as well as the individuals directing them, and Menzel was no exception. His health suffered a serious setback in 1965 from which he never fully recovered. He retired from directorship in 1966, and although he continued to teach students until 1971, as a scientist Menzel was essentially on the side-lines—occupying his mind with problems of such gravity as lunar nomenclature, and leaving it to his successors to pick up the more important pieces of the game. Death claimed him last December at the age of 75 years, after a prolonged illness; he was survived by his wife (née Florence Kreager), two daughters and six grandchildren.

Zdeněk Kopal

announcements

Meetings

16 May, **Risky ways to safety—Technological Assessment and Public Benefit**, London (Council for Science and Safety, 3/4 St Andrew's Hill, London).
19 May, **Molecular Beams in Chemistry**, London (S. Langer, The Chemical Society, Burlington House, London W1).

11–15 June, **Annual Meeting of the Society for Economic Botany**, Coral Gables, Florida (G. J. Anderson, Society for Economic Botany, Biological Sciences U43, University of Connecticut, Storrs, Conn. 06268).

31 August–2 September, Cambridge, UK, **2nd International Conference on Drag Reduction** (Secretary, 2nd Drag Reduction Conference, BHRA Fluid Engineering, Cranfield, Bedford, UK).
4–8 September, **Low Energy Ion Beams**, Salford (The Institute of Physics, 47 Belgrave Square, London SW1).

5–8 September, **Acetylenes, Allenes and Cumulenes**, Nottingham (J. F. Gibson, The Chemical Society, Burlington House, London W1).

5–9 September, **24th International Field Emission Symposium**, Oxford (G. Smith, Metallurgy and Science of Materials, Oxford University, UK).

5–9 September, **8th L. H. Gray Conference-Hypoxic Cell Sensitisers in Radiobiology and Radiotherapy**, Cambridge, UK (G. E. Adams, CRC Gray Laboratory, Mount Vernon Hospital, Northwood, Middlesex).

5–10 September, **VIIIth International Congress—International Union For the Study of Social Insects**, Wageningen, The Netherlands (International Agricultural Centre, Lawickse Allee 11, P.O. Box 88, Wageningen, The Netherlands).

5–10 September, **Biology of Connective Tissue**, Uppsala, Sweden (T. C. Laurent, Institute of Medical and Physiological Chemistry, University of Uppsala, Sweden).

6–9 September, **Plant Cell and Tissue Culture: Principles and Applications**, Columbus, Ohio (4th Annual Colloquium, College of Biological Sciences, 484 W. 12th Avenue, The Ohio State University, Columbus, Ohio 43210).

10 September, **5th European Colloquium of Geochronology, Cosmochronology, and Isotope Geology**, Pisa (G. Bonadonna, C.N.R.-Laboratorio di Geochronologia, via Cardinale Maffi, 36, 56100 Pisa, Italy).

11–16 September, **11th World Congress of Neurology**, Amsterdam (F. H. Lopes da Silva, Institute of Medical Physics, Da Costakade 45, P.O. Box 5011, Utrecht, The Netherlands).

Person to Person

Exchange furnished family house in Oxford for similar in Cape Town. Available July–September 1977. Contact Dr Donald Fraser, Department of Geology, Oxford University, England.

Exchange furnished 17th century thatched cottage 4 miles from Cambridge for house near Stanford University Medical Center, California, from February 1978 to February 1979. Car exchange also. Contact Karol Sikora, MRC Laboratory of Molecular Biology, Hills Road, Cambridge, England.

There will be no charge for this service. Send items (not more than 60 words) to Marcus Dobbs at the London office. The section will include exchanges of accommodation, personal announcements and scientific queries. We reserve the right to decline material submitted. No commercial transactions.

Reports and Publications Other countries—March

United States Department of the Interior: Geological Survey. Bulletin 1427: Surficial Geologic History of the Canyon Village Quadrangle, Yellowstone National Park, Wyoming, for use with Map 1-652. By Gerald M. Richmond. Pp. iii+35. Professional Paper 957: Lead in the Environment. Edited by T. G. Lovering. Pp. iv+90. (Washington, DC: US Government Printing Office, 1976.) [13]

Of Politics and Drug Regulation. By G. Frederick Roll. Pp. 30. (Rochester, New York: The Center for the Study of Drug Development, University of Rochester Medical Center, School of Medicine and Dentistry, 1977.) [103]

United States Department of the Interior: Geological Survey. Water-Supply Paper 2043: Water in Carbonate Rocks of the Madison Group in South-eastern Montana—a Preliminary Evaluation. By W. R. Miller. Pp. v+51+2 plates. (Washington, DC: US Government Printing Office, 1976.) [113]

American Museum Novitates, No. 2613: On the Jaw Musculature and Relationships of *Petrodromus tetradactylus* (Mammalia, Macroscelidea). By Ronn W. Coidron. Pp. 1–12. (New York: The American Museum of Natural History, 1977.) [113]

CERN—European Organization for Nuclear Research. CERN 76-24: Proceedings of the 1976 CERN School of Computing, La Grande Motte, France, 12–25 September 1976. Pp. viii+283. (Geneva: CERN, 1976.) [113]

Commission of the European Communities. Nuclear Science and Technology. Removal of Actinides from High Activity Wastes by Solvent Extraction: Outline of the Research Work at Ispra J.R.C. Laboratories. By F. Manno. Pp. 41. (Luxembourg: Office for Official Publications of the European Communities, 1976.) FB 150; £1.85. [143]

The Polychaete Worms: Definitions and Keys to the Orders, Families and Genera. By K. Fauchald. (Science Series No. 28.) Pp. 188. (Los Angeles: Natural History Museum of Los Angeles County, 1977. Published in conjunction with The Allan Hancock Foundation, University of Southern California.) \$6. [143]

Australia: Commonwealth Scientific and Industrial Research Organization. Annual Report of the Division of Protein Chemistry, 1975/1976. Pp. 39. (Melbourne: CSIRO, 1977.) [143]

Report on a Survey of the Membership of the American Astronomical Society Concerning the UFO Problem. By P. A. Sturrock. (SUIPR Report No. 681.) Pp. 202. (Stanford, California: Institute for Plasma Research, Stanford University, 1977.) [143]

CERN—European Organization for Nuclear Research. Descriptive Brochure. Pp. 20. (Geneva: CERN, 1977.) [143]

United States Department of the Interior: Geological Survey. Professional Paper 981: The Rinconada and Related Faults in the Southern Coast Ranges, California, and Their Tectonic Significance. By Thomas W. Dibblee, Jr. Pp. iv+55. (Washington, DC: US Government Printing Office, 1976.) [153]

Lei da Causalidade Universal: Teses e Hipoteses. By Waldeck Sampaio. Pp. 48. (Rio de Janeiro: Tip. Pap. Jofeleite, Ltda., Rua Santo Cristo 307, 1976.) [173]

Indian Forest Records—Silviculture. (New Series). Vol. XIII. No. 1: Population Dynamics of Forest Stands. By S. K. Seth. Pp. 179. (Delhi: Controller of Publications, 1974.) Rs. 11.50; £1.35; \$4.14. [173]

CERN—European Organization for Nuclear Research. CERN 77-03: Low-Temperature Irradiation Effects on Materials and Components for Superconducting Magnets for High-Energy Physics Applications. Edited by M. Van de Voorde. Pp. vi+69. (Geneva: CERN, 1976.) [173]

Smithsonian Contributions to Zoology, No. 243: Pseudanthessiid Copepods (Cyclopoidea). Associated with Crinoids and Echinoids (Echinodermata) in the Tropical Western Pacific Ocean. By Arthur G. Humes. Pp. iii+43. (Washington, DC: Smithsonian Institution Press, 1977. For sale by US Government Printing Office.) [173]

United States Department of the Interior: Geological Survey. Professional Paper 759-B: Physiographic Subdivisions of the Chirikov Basin, Northern Bering Sea. By D. M. Hopkins, C. H. Nelson, R. B. Perry and Tau Rho Alpha. Pp. 7+3 plates. (Washington, DC: US Government Printing Office, 1976.) [173]

United States Department of the Interior: Geological Survey. Professional Paper 959-H: Alluvial Iminite Placer Deposits, Central Virginia. By J. P. Minard, E. R. Force and G. W. Hayes. Pp. vi+15+plate 1. (Washington, DC: US Government Printing Office, 1976.) [183]

Innovation in Science Education—World Wide. By Albert V. Baez. Pp. 249. (Paris: The Unesco Press, 1976.) [223]

World Health Organization. Report of the Regional Director, July 1975 to June 1976. Pp. 102. (Copenhagen: WHO Regional Office for Europe, 1976.) [233]

Publications of the Dominion Astrophysical Observatory, Victoria, BC. Vol. XV, No. 1: Photoelectric Measures of Variable Stars Observed at Mt. Kobau (1970–73). By Graham Hill, R. W. Hilditch and E. L. Pfannenschmidt. Pp. 35. (Victoria, BC: Dominion Astrophysical Observatory, 1976.) [233]

Republic of South Africa: Department of Industries. Sea Fisheries Branch Investigational Reports. No. 111: Comparison of the Food and the Filtering Mechanism of Pilchard *Sardinops ocellata* and Anchovy *Engraulis capensis* Off South West Africa, 1971–1972. By D. P. F. King and P. R. Macleod. Pp. 29. No. 112: Exploratory Fishing for and Growth of Scallop *Pecten sulcicostatus* Off the Cape South Coast. By G. De Villiers. Pp. 23. (Cape Town: Department of Industries, Sea Fisheries Branch, 1976.) [243]

United States Department of the Interior: Geological Survey. Professional Paper 902: Precambrian Geology of the United States; An explanatory Text to Accompany the Geologic Map of the United States. By Philip B. King. Pp. iv+85. Professional Paper 965: The Relation of Geology to Stress Changes Caused by Underground Excavation in Crystalline Rocks at Idaho Springs, Colorado. By Fitzhugh T. Lee, John F. Abel, Jr., and Thomas C. Nichols. Jr. Pp. iv+47+plate 1. Professional Paper 973: The Roberts Mountains Formation, a Regional Stratigraphic Study with Emphasis on Rugose Coral Distribution. By Charles W. Merriam and Edwin H. McKee. Pp. iv+51+plates 1–12. Professional Paper 985: Biostratigraphy and Regional Relations of the Mississippian Leadville Limestone in the San Juan Mountains, Southwestern Colorado. By Augustus K. Armstrong and Bernard L. Mamet. Pp. iii+25+plates 1 and 2. Professional Paper 1000: Geological Survey Research 1976. Pp. xi+414. (Washington, DC: US Government Printing Office, 1976.) [243]

Australia: Commonwealth Scientific and Industrial Research Organization. Land Research Series, No. 38: Lands of the Alligator Rivers Area, Northern Territory. Comprising papers by R. Story, R. W. Galloway, J. R. McAlpine, J. M. Aldrick and M. A. J. Williams. Pp. 171+3 maps. (Melbourne: CSIRO, 1976.) [253]

Smithsonian Contributions to Zoology, No. 219: Phylogeny, Ontogeny, and Morphology of Living and Fossil Thaumatoctenidae (Myodocopa: Ostracoda). By Louis S. Kornicker and I. G. Sohn. Pp. iv+124. (Washington, DC: Smithsonian Institution Press, 1976. For sale by US Government Printing Office.) [253]